

**Schätzverfahren bei linearen interdependenten ökonometrischen
Systemen: Klassifizierung und Klein-Stichprobeneigenschaften
unter besonderer Berücksichtigung
von Full-Information-Verfahren**

Inaugural-Dissertation
zur Erlangung der Würde eines Doktors der Wirtschaftswissenschaften
der Wirtschaftswissenschaftlichen Fakultät
der Albert-Ludwigs-Universität
Freiburg im Breisgau

vorgelegt von
Diplom-Mathematiker Werner Scharf
aus Freiburg

1978

**Schätzverfahren bei linearen interdependenten ökonometrischen
Systemen: Klassifizierung und Klein-Stichprobeneigenschaften
unter besonderer Berücksichtigung
von Full-Information-Verfahren**

Inaugural-Dissertation
zur Erlangung der Würde eines Doktors der Wirtschaftswissenschaften
der Wirtschaftswissenschaftlichen Fakultät
der Albert-Ludwigs-Universität
Freiburg im Breisgau



vorgelegt von
Diplom-Mathematiker Werner Scharf
aus Freiburg

1978

Dekan: Prof. Dr. B. Kulp
Erstgutachter: Prof. Dr. D. Lüdeke
Zweitgutachter: Prof. Dr. G. Blümle

Die Gutachter wurden durch die Fakultät bestellt

Mündliche Prüfung: 2., 8. und 13. Februar 1978



I N H A L T

	Seite
<u>Einleitung</u>	5
Zielsetzung, Überblick, Ausblick	5
<u>I. Das Modell</u>	11
1. Vorbemerkungen	11
2. Das lineare interdependente ökonomische Modell und seine Voraussetzungen	16
3. Reduzierte Form und Identifikation	20
<u>II. Die Schätzverfahren der K-Matrix-Klasse</u>	27
1. OLS-, GLS- und IV-Methode	27
1.1. Die OLS-Schätzmethode	28
1.2. Die GLS-Schätzmethode	31
1.3. Die IV-Schätzmethode	34
2. Die Schätzverfahren der k-Klasse	36
2.1. Die ILS-Methode	37
2.2. Die 2SLS-Methode	39
1. 2SLS als IV-Schätzung	40
2. 2SLS als GLS-Schätzung	41
3. 2SLS als Ergebnis von Minimierungsprozessen	43
2.3. LISE-Methoden	45
1. Die LIML-Methode	45
2. Die LVR-Methode	54
3. Die LI/LGRV-Methode	55
4. LISE als IV-Schätzung	56

	Seite
2.4. Die k-Klasse	59
1. Herleitung und Definition der k-Klasse (IV-Interpretation der k-Klasse)	59
2. OLS, ILS, 2SLS, IISE als k-Klasse- Schätzungen	63
3. h- und Doppel-k-Klasse	65
2.5. Eigenschaften der Schätzungen der k- Klasse	67
3. Die Schätzverfahren der K-Matrix-Klasse	71
3.1. Die FI/OLS-Methode	72
3.2. Die 3SLS-Methode	75
1. 3SLS als IV-Schätzung	75
2. 3SLS als GLS-Schätzung	78
3. 3SLS als Ergebnis von Minimierungs- prozessen	79
3.3. FISS-Methoden	82
1. Die FIML-Methode	82
2. Die LGVR-Methode	88
3. Die FI/LGRV-Methode	90
4. FISS als IV-Schätzung	91
3.4. Die K-Matrix-Klasse	93
1. Herleitung und Definition der K-Matrix- Klasse (IV-Interpretation von KMC)	93
2. FI/OLS, 3SLS und FISS als K-Matrix- Klasse-Schätzungen	96
3. h-Vektor- und Doppel-K-Matrix-Klasse	97
3.5. Eigenschaften der Schätzungen der K-Matrix-Klasse	101
4. Die k-Klasse als Spezialfall der K-Matrix- Klasse	103

	Seite
5. Beziehungen zwischen verschiedenen Schätzverfahren	105
5.1. Eine Beziehung zwischen 3SLS und DKMC	106
5.2. Zusammenhänge zwischen den verschiedenen Schätzverfahren der K-Matrix-Klasse in Extremfällen	114
5.3. Schematische Zusammenfassung der wichtigsten Ergebnisse	120
<u>III. Klein-Stichprobeneigenschaften der KMC-Schätzfunktionen</u>	123
1. Möglichkeiten zur Bestimmung von Klein-Stichprobeneigenschaften von ökonometrischen Schätzfunktionen	123
2. Literaturhinweise	125
3. Nagars Verfahren zur Bestimmung des Bias und der Momentenmatrix der k-Klasse-Schätzfunktionen	129
4. Der Bias und die Momentenmatrix der K-Matrix-Klasse-Schätzfunktionen	132
5. Beweis von Theorem 4.1 und Theorem 4.2	137
5.1. Die Berechnung des Bias	139
5.2. Die Berechnung der Momentenmatrix	143
<u>Schlußbemerkungen, Zusammenfassung der Ergebnisse</u>	156
<u>Anhang</u>	
A1: Konventionen, Notationen	162
A2: Die Berechnung der Erwartungswerte	166
<u>Literaturverzeichnis</u>	175

EINLEITUNG

Zielsetzung, Überblick, Ausblick

Im Mittelpunkt des ökonometrischen Interesses vom Methodischen her stand in den letzten 30 Jahren die Schätzung von linearen interdependenten Systemen, und dies nicht zu Unrecht: So lassen sich ökonomische Beziehungen meist recht gut durch lineare Beziehungen approximieren oder in solche transformieren, und Interdependenzen, d.h. wechselseitige Abhängigkeiten zwischen ökonomischen Größen, muß man auf Grund der relativ langen Perioden zwischen zwei Beobachtungswerten annehmen. Die Schätzung der Systemparameter schließlich ist Grundlage für jede weitere Analyse (wie z.B. Prognose, Multiplikatoranalyse u.ä.).

Eine umfangreiche Schätztheorie wurde speziell für lineare interdependente Modelle entwickelt. Ihre ausführliche Darstellung findet man z.B. in den grundlegenden Lehrbüchern von Goldberger, Johnston, Christ, Malinvaud, Dhrymes, Theil und Fisk¹⁾, sowie im deutschen Sprachbereich in den

-
- 1) Goldberger, A.S.: *Econometric Theory*, New York 1964.
Johnston, J.: *Econometric Methods*, New York 1963.
Christ, C.F.: *Econometric Models and Methods*, New York/London/Sidney 1966.
Malinvaud, E.: *Statistical Methods of Econometrics*, Amsterdam 1966.
Dhrymes, P.J.: *Econometrics. Statistical Foundations and Applications*, New York/Evanston/London 1970.
Theil, H.: *Principles of Econometrics*, Amsterdam/New York 1971.
Fisk, P.R.: *Stochastically Dependent Equations*, No.21 of Griffin's Statistical Monographs and Courses, London 1967.

Werken von Lütke, Schönfeld und Schneeweiß¹⁾. Die gewissermaßen "klassische" Schätztheorie in linearen interdependenten Systemen (mit zeitlich unabhängigen Störgliedern) könnte man somit (sieht man von der Analyse der Kleinstichproben-Eigenschaften ab) als weitgehend abgeschlossen betrachten. Allerdings bleibt vom theoretischen Standpunkt her gesehen eines unbefriedigend: Viele Konzepte, die für die sogenannten Limited-information (LI-) Schätzverfahren entwickelt wurden, wurden bisher nicht auf die (effizienteren) Full-information (FI-) Schätzverfahren übertragen. Diese Lücke zu schließen, versucht die vorliegende Arbeit. Insbesondere ergibt sich daraus eine sehr einheitliche und geschlossene Klassifizierung aller Schätzverfahren und als Hauptergebnis der Arbeit eine erstmals durchgeführte Verallgemeinerung eines Verfahrens zur Berechnung von Kleinstichproben-Eigenschaften von LI-Schätzverfahren auf FI-Schätzverfahren.

Im I. Kapitel werden zunächst in den Vorbemerkungen die wichtigsten ökonometrischen Begriffe kurz erklärt, um auch Lesern ohne spezielle ökonometrische Kenntnisse die Arbeit zugänglich zu machen. Eine gewisse Vertrautheit mit den statistischen Grundbegriffen und den einfachsten Regeln der linearen Algebra und Matrizenrechnung wird jedoch vorausgesetzt²⁾. Das allgemeine lineare interdependente ökonometrische Modell und seine Voraussetzungen wird in Abschnitt 2 dargestellt. Auf das Identifikationsproblem wird nur in Hinblick auf das Verständnis der späteren Schätzproblematik eingegangen.

1) Lütke, D.: Schätzprobleme in der Ökonometrie, Würzburg 1964.

Schönfeld, P.: Methoden der Ökonometrie, Band II, München 1971.

Schneeweiß, H.: Ökonometrie, Würzburg-Wien 1971.

2) vgl. Anhang A1.

Im II. Kapitel werden alle Schätzverfahren im interdependenten Modell systematisch unter verschiedenen Gesichtspunkten dargestellt, und zwar zunächst die LI-¹⁾, dann die jeweils entsprechenden FI-Schätzverfahren. Dabei ergeben sich einige neue Aspekte und Erkenntnisse: Alle Verfahren lassen sich in einer Klasse zusammenfassen²⁾. Alle LI-Schätzverfahren haben ihr genaues FI-Pendant, aus dem sie als derjenige Spezialfall hervorgehen, bei dem in der Schätzung die kontemporären Kovarianzen zwischen verschiedenen Gleichungen des Modells nicht berücksichtigt werden. Insbesondere wird unter diesem Gesichtspunkt die Beziehung zwischen FI-Maximum-Likelihood und LI-Maximum-Likelihood-Verfahren beleuchtet. Als FI-Verallgemeinerung der gewöhnlichen Kleinst-Quadrate-(OLS-)Methode wird die Anwendung von Zellners³⁾ two-stage Aitken-estimator auf das interdependente System interpretiert. Alle Schätzverfahren können als Instrumentenvariablen-(IV-)Schätzverfahren interpretiert werden⁴⁾ (die FI-Verfahren auch als verallgemeinerte IV-Schätzungen), oder man kann sie als Ergebnis eines Minimierungsprozesses betrachten, wobei die LI-Verfahren die Spur einer Matrix, die FI-Ver-

-
- 1) Grundsätzlich werden LI-Verfahren nur als Verfahren für eine Gleichung betrachtet. Die Verallgemeinerung auf Untersysteme ist jedoch aus dem jeweiligen Zusammenhang heraus leicht ersichtlich.
 - 2) vgl. Scharf, W.: K-Matrix-Class Estimators and the Full-Information Maximum-Likelihood Estimator as a Special Case, Journal of Econometrics 4, 1976, 41-50.
 - 3) Zellner, A.: An Efficient Method of Estimating Seemingly Unrelated Regressions and Test for Aggregation Bias, Journal of the American Statistical Association 57, 1962, 348-368.
 - 4) vgl. Tabelle 1b) am Ende von Kapitel II.

fahren die Determinante derselben Matrix als "Maß" verwenden¹⁾. FI-Verallgemeinerungen von k-, h- und Doppel-k-Klasse werden vorgeschlagen. Die Beziehungen zwischen den einzelnen Klassen von Schätzverfahren sind in Schema ²⁾ als Diagramm dargestellt. Schließlich wird in Abschnitt 5.1 eine bekannte Beziehung zwischen Doppel-k-Klasse und zweistufigem Kleinst-Quadrate-(2SLS-)Verfahren auf den entsprechenden Full-Information Fall verallgemeinert. Neben diesen neuen Aspekten der Schätztheorie konnten die eigentlichen Schätzverfahren relativ knapp (nicht immer mit letzter mathematischer Strenge) dargestellt werden, einerseits um den Blick für die größeren Zusammenhänge nicht zu verbauen, zum anderen, da in den Büchern von Goldberger, Schönfeld und Dhrymes³⁾, an deren Darstellungen ich mich im wesentlichen angelehnt habe, schon detaillierte Beschreibungen der eigentlichen Schätzverfahren existieren.

Als eigentliches Hauptresultat der vorliegenden Arbeit ist jedoch das in Theorem 4.1 und 4.2 zusammengefaßte Ergebnis von Kapitel III zu betrachten: Erstmals wird Nagars Verfahren⁴⁾ zur approximativen Berechnung von Kleinstichproben-Eigenschaften der Limited-Information-Schätzverfahren (speziell 2SLS) auf Full-Information-Verfahren (speziell 3SLS) verallgemeinert. Dieses Ergebnis stellt

1) vgl. Tabelle 1a) am Ende von Kapitel II.

2) vgl. Schema 1 am Ende von Kapitel II.

3) Goldberger, A.S.: Econometric Theory, a.a.O.Kap.7.

Schönfeld, P.: Methoden der Ökonometrie, a.a.O.Kap. 16-18.

Dhrymes, P.J.: Econometrics, a.a.O. Kap. 4-7.

4) Nagar, A.L.: The Bias and Moment Matrix of the General k-Class Estimators of the Parameters in Simultaneous Equations, Econometrica 27, 1959, 575-595.

einen der sehr wenigen Beiträge zu den bislang sehr geringen Kenntnissen über Kleinstichproben-Eigenschaften von Full-Information-Schätzverfahren dar.

Dem oben beschriebenen speziellen Anliegen der Arbeit entsprechend, wurde auf die Berücksichtigung einiger in den letzten Jahren in den Vordergrund tretenden theoretischen Methoden der ökonometrischen Analyse verzichtet, wie z.B. den Zeitreihenaspekt (hier besonders die Spektralanalyse¹⁾), die Laganalyse²⁾ oder den Standpunkt der Bayesschen Statistik³⁾. Die stärkere Einbeziehung dieser Aspekte in die Schätztheorie (besonders bei allgemeineren Annahmen über die Störglieder) ist eine Zukunftsaufgabe. Ob sich die Schätzverfahren dann noch ähnlich einheitlich darstellen und klassifizieren lassen, wie in der vorliegenden Arbeit gezeigt wird, ist eine offene Frage.

Schließlich sei noch erwähnt, daß numerische Anwendungen der verschiedenen Schätzverfahren an Hand von kleineren Demonstrationsmodellen in den meisten Lehrbüchern enthalten sind. Besonders was die numerisch äußerst schwierig zu handhabenden Schätzungen nach der Maximum-Likelihood-Methode betrifft, war hier die Arbeit von Lüdeke⁴⁾ bahnbrechend. Die speziell bei der praktischen Anwendung der Schätzverfahren auf ein größeres Modell⁵⁾ auftretenden

1) Dhrymes, P.J.: Econometrics, a.a.O. Kap.9-12.

2) Dhrymes, P.J.: Distributed Lags, San Francisco 1971.

3) Zellner, A.: An Introduction to Bayesian Inference in Econometrics, New York 1971.

4) Lüdeke, D.: Schätzprobleme in der Ökonometrie, a.a.O. vgl. insbes. Abschn. 3.5.2.2.

5) Lüdeke, D.: Ein ökonometrisches Vierteljahresmodell für die Bundesrepublik Deutschland, Tübingen 1969.

Probleme, die man keinesfalls unterschätzen sollte, wurden ausführlich von Friedrich¹⁾ behandelt. Auf Grund dessen konnte ich in der vorliegenden Arbeit auf eine numerische Demonstration der einzelnen Schätzverfahren verzichten und mich, dem rein theoretischen Charakter der Arbeit entsprechend, auf die in den vorangegangenen Abschnitten beschriebene Problematik beschränken.

An dieser Stelle verbleibt mir die angenehme Pflicht, meinem verehrten Lehrer, Herrn Professor Dr. D. Lüdke, für seine zahlreichen Anregungen und das Interesse, das er jederzeit meiner Arbeit entgegenbrachte, aufrichtig zu danken.

1) Friedrich, D.: Schätzprobleme sowie Verfahren und Möglichkeiten der Analyse ökonometrischer Modelle, Tübingen 1974.

I. DAS MODELL

1. Vorbemerkungen

Um reale ökonomische Vorgänge, Abhängigkeiten oder Interaktionen nicht nur beobachten zu können, sondern um sie auch theoretisch analysieren und daraus allgemeinere Zusammenhänge erkennen zu können, bedient man sich der "Sprache" der Mathematik, die es erlaubt, auch komplexe Vorgänge kurz, prägnant und allgemeingültig auszudrücken. Zudem stellt sie Mittel zur Verfügung, mit Hilfe derer man aus den gewonnenen theoretischen Erkenntnissen Folgerungen ziehen kann, die man letztlich dann wieder auf den realen Wirtschaftsvorgang überträgt.

Als theoretisches Modell bezeichnet man somit die in die "Sprache" der Mathematik übersetzte wirtschaftliche Realität!¹⁾ Die Übersetzung erfolgt wie bei gewöhnlichen Sprachen mit Hilfe von Begriffsidentifikationen. So können z.B. ökonomische Größen mit Variablen, ökonomische Vorgänge mit Funktionen identifiziert werden. Die ökonomische Größe "privater Verbrauch" wird dann z.B. mit der Variablen C gleichgesetzt, die ökonomische Feststellung, daß das verfügbare Einkommen der privaten Haushalte den privaten Konsum beeinflusst, wird "übersetzt" in die Funktion $C=f(Y)$. Hier zeigt sich nun allerdings eine Schwierigkeit: In der ökonomischen Realität sind Abhängigkeiten nicht unmittelbar beobachtbar, sondern nur die eigentlichen ökonomischen Größen innerhalb eines Zeitraums. So könnte es z.B. sein, daß die Funktion $C=f(Y)$ nicht die tatsächlichen ökonomischen Abhängigkeiten voll abdeckt, da vielleicht noch andere Größen den privaten Konsum beeinflussen. Ein Ausweg aus dieser Schwierigkeit sind die von der mathematischen Statistik entwickelten Verfahren, mit Hilfe derer es möglich ist, anhand der beobachtbaren

1) Der Modellbegriff ist hier also insofern recht eng gefaßt, als er speziell die mathematische Darstellung beinhaltet.

ökonomischen Größen auch Rückschlüsse auf die Gültigkeit von Funktionszusammenhängen wie im obigen Beispiel zu ziehen.

Wie bereits erwähnt, werden gewöhnlich ökonomische Größen als Variable, ökonomische Abhängigkeiten als Funktionen dargestellt (obwohl dies natürlich nicht die einzige Möglichkeit der Darstellungsweise ist). Die Frage, die sich dabei aufwirft, ist die nach dem Funktionstyp: Der einfachste mathematische Funktionstyp ist der einer *l i n e a - r e n* Funktion. Er hat den großen Vorteil, daß er analytisch am einfachsten zu handhaben ist (für ihn ist die am weitest gehende mathematische Theorie - die lineare Algebra - entwickelt), er eine große Reihe von Zusammenhängen sowieso schon abdeckt und außerdem sich viele kompliziertere Funktionstypen durch einfache Transformationen in ihn überführen lassen. Zwar können unter Umständen ökonomische Vorgänge durch kompliziertere Funktionstypen besser beschrieben werden. Dieser vermeintliche Vorteil wird jedoch i.a. dadurch zunichte gemacht, daß man für nichtlineare Funktionen diese bessere Beschreibung der Realität nicht nachweisen kann, da für kompliziertere Funktionen keine annähernd so weit entwickelte Schätztheorie existiert wie die im II. Kapitel beschriebene für lineare Funktionen. Aus den oben genannten Gründen werden in dieser Arbeit nur lineare Funktionen behandelt.

Ein theoretisches Modell zur Beschreibung von realem ökonomischen Geschehen - kurz *ö k o n o m i s c h e s* *M o d e l l* - besteht i.a. nicht nur aus einer Funktion bzw. Gleichung, sondern - der Vielfalt des ökonomischen Geschehens entsprechend - aus einer ganzen Reihe von Gleichungen, die nicht isoliert betrachtet werden können. Ein (sehr einfaches) lineares ökonomisches Modell zur Beschreibung einer Volkswirtschaft:

$$(1.1) \quad \begin{aligned} C &= a + bY \\ Y &= C + I \end{aligned}$$

illustriert dies. C bedeutet hier Konsum, Y Volkseinkommen, I Investitionen. Wie man sieht, ist einerseits C eine Funktion von Y , andererseits aber auch Y über eine Definitionsgleichung eine Funktion von C . Diese gegenseitige Abhängigkeit ökonomischer Variablen bezeichnet man als *Interdependenz*. Ein ökonomisches Modell, das nur aus linearen Gleichungen besteht und bei dem Interdependenzen der eben beschriebenen Art auftreten, heißt *lineares interdependentes Modell*. Die Variablen eines linearen interdependenten Modells lassen sich in der folgenden Weise klassifizieren: Alle Variablen, die über das Modell erklärt werden, die also auf der linken Seite einer Gleichung stehen (in (1.1) also C und Y) heißen *endogene Variable*. Die übrigen Variablen, von denen man annimmt, daß sie außerhalb des Modells bestimmt wurden, heißen *exogen* (in (1.1) also I).

Nun genügen jedoch funktionelle Zusammenhänge zwischen ökonomischen Variablen wie in Beispiel (1.1) nicht, um die ökonomische Realität ausreichend exakt zu beschreiben. Ein wichtiger Einflußfaktor, nämlich die Zeit, muß noch hinzukommen. So besagt der in der ersten Gleichung von (1.1) ausgedrückte lineare Einfluß von Y auf C zunächst nicht viel, wenn man nicht weiß, zu welchem Zeitpunkt er stattfindet. Da ökonomische Größen nur zu bestimmten diskreten Zeitpunkten erhoben werden (d.h. den Variablen ein Wert zugeordnet wird), kann eine Gleichung der Form $C = a + bY$ vielerlei bedeuten: C kann eine Funktion des jeweils zum selben Zeitpunkt erhobenen Y oder des jeweils zu einem der vielen früheren Zeitpunkte erhobenen Y sein. Dieser Sachverhalt wird ausgedrückt durch folgende Schreibweise:

$$C_t = a + bY_t \quad \text{bzw.} \quad C_t = a + bY_{t-1} \quad \text{bzw.} \quad C_t = a + bY_{t-2} \quad \text{usw.}$$

Der Zeitraum von $t-1$ bis t ist jeweils der Zeitraum (die Periode) von einem Erhebungszeitpunkt bis zum nächsten. Er wird i.a. konstant sein und ist z.B. ein Jahr, ein Quartal oder ein Monat. Das Beispielmmodell (1.1) wird man nun präzisieren müssen in der Form

$$(1.2) \quad \begin{aligned} C_t &= a + bY_{t-1} \\ Y_t &= C_t + I_t \end{aligned}$$

wenn man Anlaß zu der Vermutung hat, daß eher das Einkommen der Vorperiode ("verzögert") den Konsum der laufenden Periode beeinflußt, als das Einkommen derselben Periode ("unverzögert"). Modell (1.2) induziert eine andere, für die statistische Analyse wichtigere Einteilung der ökonomischen Variablen: die Einteilung in *g e m e i n s a m a b h ä n g i g e* und *v o r h e r b e s t i m m t e* Variablen. Gemeinsam abhängige Variablen sind alle unverzögerten endogenen Variablen, vorherbestimmte Variablen sind die verzögerten endogenen und alle exogenen Variablen. Modelle mit verzögerten endogenen Variablen heißen *d y n a m i s c h*. Die Verzögerung als solche bezeichnet man auch als *L a g*.

Das bis hierher beschriebene ökonomische Modell entspricht Modellen wie sie in der Wirtschaftstheorie verwendet werden. Die Ökonometrie geht in der Analyse von ökonomischen Modellen noch weiter: Man begnügt sich nicht damit, einen allgemeinen linearen Zusammenhang der Art wie in (1.2) als Hypothese zu postulieren. Man möchte vielmehr wissen, welche konkreten Werte die Koeffizienten oder *P a r a m e t e r* (z.B. a, b in (1.2)) der Gleichungen eines Modells besitzen. Die mathematische Statistik bietet Möglichkeiten, an Hand von *B e o b a c h t u n g e n* (das sind die zu verschiedenen Zeitpunkten erho-

benen Werte) von ökonomischen Variablen Rückschlüsse auf die Größe der Parameter zu machen. Dieser Vorgang heißt **Parameterschätzung** und wird in Kapitel II für das lineare interdependente Modell ausführlich behandelt. Gleichzeitig ermöglicht es die mathematische Statistik, Aussagen darüber zu treffen, ob ein hypothetisch angenommener Zusammenhang zwischen ökonomischen Größen auch wirklich besteht. Dieser Aspekt der Ökonometrie wird in der vorliegenden Arbeit nicht berücksichtigt, d.h. es wird im folgenden angenommen, daß die vorgegebenen Zusammenhänge tatsächlich existieren.

Ein weiteres Konzept der Ökonometrie ist die **Stochastisierung** eines ökonomischen Modells, die erst die oben genannte Parameterschätzung möglich macht. D.h. man nimmt realistischerweise an, daß nicht alle Einflüsse, die auf eine ökonomische Größe wirken, erfaßt oder modellmäßig dargestellt werden können. Daher führt man ein **Störglied** als Zufallsvariable ein, das alle übrigen Einflüsse beinhaltet. Die Konsumgleichung in Beispiel (1.2) würde demnach in stochastisierter Form wie folgt aussehen:

$$C_t = a + bY_{t-1} + u_t$$

Die Zufallsvariable u_t folgt also einer Wahrscheinlichkeitsverteilung (meist wird Normalverteilung angenommen). Da C_t linear von u_t abhängt, ist mit u_t auch C_t eine Zufallsvariable. Man betrachtet alle Beobachtungen von C_t (wie natürlich jeder anderen endogenen Variablen) als Stichprobe (d.h. als Realisierung einer Zufallsvariablen) vom Stichprobenumfang T (wenn man T Beobachtungen: $t=1, \dots, T$ hat) aus einer hypothetischen Grundgesamtheit. Exogene Variable werden oft als nichtstochastisch angenommen.

Ein stochastisiertes ökonomisches Modell, dessen Parameter bestimmte Werte annehmen können, heißt kurz **ö k o - n o m e t r i s c h e s M o d e l l**. Ein ökonometrisches Modell versehen mit bestimmten Parameterwerten und einer bestimmten Verteilung für die Störglieder heißt **S t r u k - t u r** des Modells.

Bei größeren ökonometrischen Modellen bietet sich als Kurzschreibweise die Matrixschreibweise an: Die Beobachtungen von gemeinsam abhängigen Variablen und vorherbestimmten Variablen werden jeweils in einer Matrix zusammengefaßt, und zwar so, daß jede Spalte der Matrix den gesamten Beobachtungssatz einer Variablen bildet. Ebenso werden die Parameter und die in jeder Gleichung auftretenden Störglieder jeweils zu einer Matrix zusammengefaßt. Die Koeffizienten der abhängigen Variablen sind nur da a priori gleich eins, wo die die abhängigen Variablen auf der linken Seite einer Gleichung stehen.

Nach diesen Vorbemerkungen kann man nun das allgemeine lineare interdependente ökonometrische Modell in der im nächsten Abschnitt aufgezeigten Art definieren.

2. Das lineare interdependente ökonometrische Modell und seine Voraussetzungen.

Es wird in der vorliegenden Arbeit folgendes lineare interdependente Modell von N Gleichungen betrachtet:

$$(2.1) \quad Y\Gamma + XB + U = 0$$

Es ist dabei

Y : die TxN-Matrix der Beobachtungen der gemeinsam abhängigen Variablen

Γ : die NxN-Matrix der Koeffizienten der gemeinsam abhängigen Variablen

X : die $T \times R$ -Matrix der Beobachtungen der vorherbestimmten Variablen

B : die $R \times N$ -Matrix der Koeffizienten der vorherbestimmten Variablen

U : die $T \times N$ -Matrix der Störglieder.

T ist die Zahl der Beobachtungen.

Noch kompakter läßt sich Modell ¹⁾ (2.1) schreiben als

$$(2.1') \quad ZA + U = 0$$

$$\text{wo } Z := (Y, X), \quad A := \begin{pmatrix} \Gamma \\ B \end{pmatrix}.$$

Γ ist dadurch normiert, daß die Diagonalelemente γ_{ii} ($i=1, \dots, N$) alle $=1$ gesetzt sind. Außerdem wird vorausgesetzt, daß Γ nichtsingulär ist, die Inverse also existiert.

Für den Rang von X gelte $\text{rg}X = R \leq T$.

Die Koeffizientenmatrizen Γ und B müssen nicht vollbesetzt sein. Es werden im Gegenteil in der Regel einige Elemente (natürlich nicht die Diagonalelemente von Γ) von Γ und B a priori als Null vorausgesetzt sein (auf Grund des noch später zu behandelnden Identifikationsproblems). Diese a priori Nullrestriktionen besagen nichts anderes, als daß die zu dem als Null vorausgesetzten Koeffizienten gehörige Variable zwar im gesamten Modell, nicht jedoch in der entsprechenden Gleichung auftritt.

Unter Berücksichtigung von a priori Nullrestriktionen kann man die i -te Gleichung des Modells (2.1) schreiben als

$$(2.2) \quad y_i = Y_i \gamma_i + X_i \beta_i + u_i$$

1) Hier und im Folgenden wird immer vorausgesetzt, daß Definitionsgleichungen aus dem Modell eliminiert wurden. Zu einer ausführlichen Darstellung der durch Definitionsgleichungen in der Praxis auftretenden Problematik siehe Friedrich, D.: a.a.O., Kapitel I, insbesondere Abschnitt 3.2.

mit

y_i : der $T \times 1$ -Vektor der Beobachtungen der abhängigen Variablen der i -ten Gleichung

Y_i : die $T \times n_i$ -Matrix der Beobachtungen der in der i -ten Gleichung rechtsseitig auftretenden erklärenden abhängigen Variablen

γ_i : der $n_i \times 1$ -Vektor der Koeffizienten von Y_i

X_i : die $T \times r_i$ -Matrix der Beobachtungen der in der i -ten Gleichung auftretenden vorherbestimmten Variablen

β_i : der $r_i \times 1$ -Vektor der Koeffizienten von X_i

u_i : der $T \times 1$ -Vektor des Störgliedes der i -ten Gleichung

(y_i, Y_i) , X_i , γ_i , β_i sind also jeweils die Untermatrizen bzw. -vektoren von Y , X , Γ , B , die unter Beachtung von Nullrestriktionen der i -ten Gleichung entsprechen.

Die Matrix U der Störglieder sei auf zweierlei Weise aufgeteilt, einmal spaltenweise: $U = (u_1, \dots, u_N)$, zum andern zeilenweise: $U = (u(1), \dots, u(t), \dots, u(T))'$, wo $u(t)$ ein $N \times 1$ -Vektor ist, der aus den Elementen der t -ten Zeile von U besteht.

In Kurzform kann man die i -te Gleichung auch schreiben als

$$(2.2') \quad y_i = Z_i \alpha_i + u_i \quad \text{mit } Z_i := (Y_i, X_i), \quad \alpha_i := \begin{pmatrix} \gamma_i \\ \beta_i \end{pmatrix}.$$

Es sei an dieser Stelle betont, daß im allgemeinen Modell (2.1) bzw. (2.2) stets auch der einfachste Fall eines Modells mit nur einer Gleichung enthalten ist: Man muß nur $N=1$ setzen und erhält

$$y = X\beta + u,$$

eine Gleichung mit R vorherbestimmten Variablen.

Die folgenden stochastischen Annahmen¹⁾ über das Modell werden gemacht:

1. Die $u(t)$ bilden einen N-dimensionalen stationären stochastischen Prozeß mit
 - (a) $E\{u(t)\} = 0$ für alle t
 - (b) $E\{u(t)u(t)'\} = \Sigma$ für alle t , mit positiv definitem Σ
 - (c) $E\{u(t)u(t')'\} = 0$ für alle $t \neq t'$
 - (d) Die $u(t)$ sind stochastisch unabhängig und identisch verteilt.
2. Die $x(t)$ (die spaltenweise geschriebenen Reihen von X) bilden einen stationären stochastischen Prozeß (sofern sie nicht nichtstochastisch sind) mit
 - (a) $\text{plim } T^{-1}X'X = M$ existiert und ist positiv definit
 - (b) Es gilt $\text{plim } T^{-1}X'U = 0$
 - (c) Die Absolutwerte der Elemente von X sind für $T \rightarrow \infty$ beschränkt.
3. Das Modell ist stabil, falls es dynamisch ist.

Die $N \times N$ -Kovarianzmatrix Σ besteht aus den Elementen

$\sigma_{ih} = E\{u_i(t)u_h(t)\}$ ($i, h=1, \dots, N$). Es ist

$\sigma_i := \begin{pmatrix} \sigma_{i1} \\ \vdots \\ \sigma_{Ni} \end{pmatrix}$ die i -te Spalte von Σ .

Eine wichtige Folgerung aus den obigen Annahmen ist hier ohne Beweis angegeben:

Folgerung 1: Unter den obigen stochastischen Annahmen gilt
 $\text{plim } T^{-1}U'U = \Sigma$

1) Die stochastischen Voraussetzungen sind auf die Bedürfnisse der vorliegenden Arbeit beschränkt. Eine mathematisch strengere Darstellung der stochastischen Annahmen findet man bei Schönfeld, P.: Methoden der Ökonometrie, Bd.II, a.a.O.S.158ff..Dort ist auch der Begriff des stochastischen Prozesses erklärt.

Zu bemerken ist noch, daß die Annahmen 1(b) und (c) in Kroneckerprodukt-Schreibweise zusammengefaßt werden können zu

$$1(bc) \quad E\{uu'\} = I \otimes \Sigma$$

$$\text{wo } u = \begin{pmatrix} u(1) \\ \vdots \\ u(T) \end{pmatrix} \text{ und } I \text{ die } T \times T\text{-Einheitsmatrix ist,}$$

oder

$$1(bc)' \quad E\{u^* u^{*'}\} = \Sigma \otimes I$$

$$\text{wo } u^* = \begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix}.$$

3. Reduzierte Form und Identifikation

Die Darstellungsweise (2.1) des allgemeinen linearen interdependenten ökonomischen Modells heißt **S t r u k t u r f o r m** des Modells und ist die Darstellungsweise, die die direkten Einflüsse der erklärenden Variablen auf eine bestimmte ökonomische Größe, nämlich die abhängige Variable z.B. der i-ten Gleichung, wiedergibt. Sie ist insofern die für die Modellspezifizierung und die Schätzung relevante Form. Eine weitere Darstellungsweise des Modells ist die **r e d u z i e r t e F o r m**, die den (indirekten)Einfluß nur der vorherbestimmten Variablen des Modells auf die Abhängigen wiedergibt. In einem statischen Modell (also einem Modell ohne verzögerte endogene Variablen) fällt die reduzierte Form mit der sogenannten **F i n a l f o r m** zusammen, die bei dynamischen Modellen den Einfluß der exogenen und der Anfangswerte der endogenen Variablen auf die abhängigen Variablen widerspiegelt.

Von der Strukturform (2.1) kommt man zur reduzierten Form sehr leicht. Da Γ als nichtsingulär vorausgesetzt ist,

muß man Gleichung (2.1) nur mit Γ^{-1} postmultiplizieren. Daraus ergibt sich folgende Definition der reduzierten Form:

Die reduzierte Form von Modell (2.1) lautet

$$(3.1) \quad Y = X\Pi + V \quad \text{mit } \Pi = -B\Gamma^{-1}, V = -U\Gamma^{-1}$$

Die i-te Gleichung des Gleichungssystems (3.1) lautet dann

$$(3.2) \quad y_i = X\pi_i + v_i$$

wo y_i , π_i , v_i jeweils die i-ten Spalten von Y , Π und V sind.

Da V durch eine lineare Transformation aus U hervorgeht, lassen sich aus den Annahmen 1 in Abschnitt 2 leicht die stochastischen Eigenschaften von $v(t)$, den spaltenweise geschriebenen Zeilenvektoren von V , bestimmen:

- (a) $E\{v(t)\} = -\Gamma^{-1}E\{u(t)\} = 0$
- (b) $E\{v(t)v(t)'\} = \Gamma^{-1}E\{u(t)u(t)'\}\Gamma^{-1} = \Gamma^{-1}\Sigma\Gamma^{-1} = \Omega > 0$
- (c) $E\{v(t)v(t')'\} = \Gamma^{-1}E\{u(t)u(t')'\}\Gamma^{-1} = 0 \quad \text{für } t \neq t'$
- (d) Die $v(t)$ sind stochastisch unabhängig und identisch verteilt.

Wie wir gesehen haben, läßt sich von der Strukturform eines Modells eindeutig auf seine reduzierte Form schließen (nämlich durch die Lineartransformation Γ^{-1}). Läßt sich nun auch umgekehrt von einer vorgegebenen reduzierten-Form-Gleichung eindeutig auf die Strukturform schließen? Daß dies im allgemeinen nicht der Fall ist, zeigt folgender Satz:

Satz 3.1: Zwei Strukturformen, die durch die Parameter (Γ, B) bzw. (Γ^*, B^*) gekennzeichnet sind, besitzen genau dann dieselbe reduzierte Form, wenn sie sich durch eine nichtsinguläre Lineartransformation ineinander überführen lassen.

Beweis: 1. Es seien (Γ, B) und (Γ^*, B^*) die beiden Strukturen, die durch eine nichtsinguläre Lineartransformation verbunden sind:

$$\Gamma = \Gamma^* \Lambda, \quad B = B^* \Lambda, \quad U = U^* \Lambda \quad (|\Lambda| \neq 0)$$

Dann folgt

$$\Pi = -B\Gamma^{-1} = -B^* \Lambda \Lambda^{-1} \Gamma^{*-1} = -B^* \Gamma^{*-1} = \Pi^*$$

$$V = -U\Gamma^{-1} = -U^* \Lambda \Lambda^{-1} \Gamma^{*-1} = -U^* \Gamma^{*-1} = V^*, \text{ damit auch } \Omega = \Omega^*.$$

Also sind die reduzierten Formen gleich.

2. Sei umgekehrt $Y = X\Pi + V$ die reduzierte Form für die beiden Strukturen, die durch (Γ, B) bzw. (Γ^*, B^*) gekennzeichnet sind. Es ist dann

$$-B\Gamma^{-1} = \Pi = -B^* \Gamma^{*-1}$$

$\Rightarrow B = B^* \Gamma^{*-1} \Gamma = B^* \Lambda$ mit nichtsingulärem Λ , da Γ^*, Γ nichtsing. und $\Gamma = \Gamma^* \Gamma^{*-1} \Gamma = \Gamma^* \Lambda$ mit demselben nichtsingulären Λ .

Außerdem ist $-U\Gamma^{-1} = V = -U^* \Gamma^{*-1}$

$\Rightarrow U = U^* \Gamma^{*-1} \Gamma = U^* \Lambda$ mit demselben nichtsingulären Λ .

q.e.d.

Aus Satz 3.1 erkennt man, daß eine eindeutige Beziehung zwischen Strukturform und reduzierter Form eines Modells genau dann besteht, wenn jede zulässige nichtsinguläre Lineartransformation die gegebene Strukturform bis auf Normierung in sich selbst überführt. Zulässig ist eine Transformation, wenn sie etwaige vorgegebene Parameterrestriktionen nicht verändert.

Was sind nun die Bedingungen dafür, daß man von der reduzierten Form eines Modells eindeutig auf die Parameter einer bestimmten Gleichung der Strukturform schließen kann?

Seien γ und β die Parameter der i-ten Gleichung. Wegen $-\Pi\Gamma = B$ gilt dann offensichtlich $-\Pi\gamma = \beta$ bzw.

$$(3.3) \quad (\Pi, I) \begin{pmatrix} \gamma \\ \beta \end{pmatrix} = 0$$

(3.3) bildet ein lineares homogenes Gleichungssystem von R Gleichungen und $R+N$ Unbekannten. Seine Koeffizientenmatrix (π, I) besitzt, da I nichtsingulär ist, den Rang R . Dies bedeutet jedoch, daß (bis auf Normierung) $R+N-R=N$ verschiedene Lösungen existieren. D.h. Gleichungssystem (3.3) ist nicht eindeutig lösbar (sofern $N>1$), und man kann in diesem allgemeinen Fall nicht eindeutig von der reduzierten Form auf die Koeffizienten der i -ten Strukturgleichung schließen. Es seien nun jedoch gewisse Parameter aus γ und β gleich Null und $\bar{\gamma}_i, \beta_i$ seien die Parameter der i -ten Gleichung, die ungleich Null sind (ohne Normierung, daher der Querstrich über $\bar{\gamma}_i$). Bei entsprechender Umordnung der Koeffizienten und Aufspaltung der Koeffizientenmatrizen wird Gleichungssystem (3.3) zu

$$\begin{pmatrix} \pi_{11} & \pi_{12} & I_1 & 0 \\ \pi_{21} & \pi_{22} & 0 & I_2 \end{pmatrix} \begin{pmatrix} \bar{\gamma}_i \\ 0 \\ \beta_i \\ 0 \end{pmatrix} = 0, \text{ bzw.} \\ (3.3') \quad \begin{pmatrix} \pi_{11} & I_1 \\ \pi_{21} & 0 \end{pmatrix} \begin{pmatrix} \bar{\gamma}_i \\ \beta_i \end{pmatrix} = 0$$

Gleichungssystem (3.3') ist ein homogenes Gleichungssystem von $R=R_1+\bar{R}_1$ Gleichungen¹⁾ und N_1+R_1 Unbekannten. Eine bis auf Normierung eindeutige Lösung von (3.3') existiert genau dann, wenn der Rang der in (3.3') links stehenden Matrix gleich N_1+R_1-1 ist. Eine notwendige Bedingung hierfür ist natürlich, daß die Zeilenanzahl R der Koeffizientenmatrix von (3.3') mindestens gleich N_1+R_1-1 ist, daß also gilt $\bar{R}_1 \geq N_1-1$. Die oben angegebene Rangbedingung

1) N_1 ist die Anzahl der in der i -ten Gleichung insgesamt vorkommenden gemeinsam abhängigen Variablen (einschl. der linksseitig Abhängigen), n_i die Anzahl der erklärenden abhängigen Variablen der i -ten Gleichung. R_i und r_i werden äquivalent für die Anzahl der in der i -ten Gleichung vorkommenden vorherbestimmten Variablen gebraucht.

für die Koeffizientenmatrix ist (wegen $\text{rg } I_1 = R_1$) äquivalent mit der Bedingung $\text{rg } \Pi_{21} = N_1 - 1$.

Zusammenfassend läßt sich folgender Satz formulieren:
Satz 3.2: Von den Parametern der reduzierten Form eines Modells kann man genau dann (bis auf Normierung eindeutig) auf die Parameter einer (z.B. der i-ten) Strukturgleichung schließen, wenn die $\bar{R}_i \times N_i$ -Untermatrix Π_{21} von Π den Rang $N_i - 1$, also die um 1 verminderte Anzahl der in der i-ten Gleichung auftretenden abhängigen Variablen, hat. Eine notwendige Bedingung hierfür ist, daß $\bar{R}_i \geq N_i - 1$ ist, die Anzahl der aus der i-ten Gleichung ausgeschlossenen vorherbestimmten Variablen also mindestens so groß wie die um 1 verminderte Anzahl der in der i-ten Gleichung auftretenden abhängigen Variablen ist.

Die Frage, ob man von einer Reduzierten-Form-Gleichung eindeutig auf die Strukturform zurückschließen kann, ist eng verknüpft mit der wichtigen Frage der **I d e n t i f i z i e r b a r k e i t** einer Struktur. Man kann nämlich zeigen¹⁾, daß zwei verschiedene Strukturen dann und nur dann dieselbe bedingte Verteilungsfunktion von Y (das ja als Zufallsvariable eine Wahrscheinlichkeitsverteilung besitzt) induzieren, wenn sie durch eine nichtsinguläre Lineartransformation verbunden sind. Die beiden Strukturen sind dann **b e o b a c h t u n g s ä q u i v a l e n t**, da beide mit der Verteilung von Y verträglich sind, man also von einer Stichprobe von Y (d.h. von den Beobachtungswerten von Y) nicht mehr eindeutig auf die Parameter einer bestimmten Struktur schließen kann. Eine Struktur heißt nun genau dann identifizierbar, wenn jede zu ihr beobachtungsäquivalente Struktur (also jede durch eine zulässige

1) vgl. Lüdtke, D.: Schätzprobleme in der Ökonometrie, a.a.O. S.31ff.

nichtsinguläre Lineartransformation aus ihr hervorgegangene Struktur) bis auf einen Normierungsfaktor mit ihr übereinstimmt.

Aus dem soeben Gesagten in Verbindung mit Satz 3.1 kann man folgenden Satz formulieren:

Satz 3.3: Die Koeffizienten der Strukturform eines Modells sind genau dann identifizierbar, wenn man (bis auf Normierung) eindeutig von ihrer reduzierten Form auf sie zurückschließen kann. Eine Strukturgleichung ist genau dann identifizierbar, wenn man von den Koeffizienten der reduzierten Form eindeutig (bis auf Normierung) auf die in dieser Gleichung vorkommenden Koeffizienten schließen kann.

Satz 3.3 zeigt, daß Satz 3.2 gleichzeitig Aussagen über die Identifizierbarkeit einer Strukturgleichung macht. Im Zusammenhang mit der Identifizierbarkeit nennt man dabei die Bedingung $rg \Pi_{21} = N_1 - 1$ die Rangbedingung, die Bedingung $\bar{R}_1 \geq N_1 - 1$ die Abzählbedingung für die Identifizierbarkeit der i-ten Strukturgleichung. Die Abzählbedingung gewährleistet allerdings noch nicht die Identifizierbarkeit, da sie nur eine notwendige, nicht jedoch hinreichende Bedingung darstellt. Auf Grund der Abzählbedingung unterscheidet man folgende Fälle:

Definition 3.1: Die i-te Strukturgleichung eines Modells heißt nicht identifizierbar, wenn $\bar{R}_1 < N_1 - 1$. Sie heißt genau identifizierbar, wenn sie identifizierbar ist und $\bar{R}_1 = N_1 - 1$ gilt. Sie heißt überidentifizierbar, wenn sie identifizierbar ist und $\bar{R}_1 > N_1 - 1$ gilt.

Satz 3.2 gibt die Hauptidegebnisse über Identifizierbarkeit und die Eineindeutigkeit der Beziehungen zwischen reduzierter Form und Strukturform wieder, soweit sie in

der vorliegenden Arbeit zum Verständnis des Folgenden benötigt werden. Es sei betont, daß das Problem der Identifizierbarkeit hier nur in einem sehr speziellen Zusammenhang bei speziellen Parameterrestriktionen (nämlich von Nullrestriktionen) behandelt wurde. Identifizierbarkeitsbedingungen für allgemeine lineare Restriktionen sind bei Lüdeke¹⁾ angegeben, eine ausführliche theoretische Darstellung des Identifikationsproblems bei Schönfeld²⁾. Dort sind auch weitere Literaturhinweise zu finden.

1) Lüdeke, D.: Schätzprobleme in der Ökonometrie, a.a.O. S.48ff.

2) Schönfeld, P.: Methoden der Ökonometrie, Band II, a.a.O. S.185ff.

II. DIE SCHÄTZVERFAHREN DER K-MATRIX-KLASSE

In diesem Kapitel werden die wichtigsten Schätzverfahren zur Schätzung der Parameter eines linearen interdependenten Modells unter den Modellannahmen des Kapitels I, Abschnitt 2 und unter Beschränkung auf Nullrestriktionen der Strukturparameter dargestellt. Wo zur Herleitung von Ergebnissen weitere Annahmen als die von Kapitel I nötig sind, wird dies an der jeweiligen Stelle erwähnt.

Die Schätzverfahren werden unterteilt in Schätzverfahren für eine Strukturgleichung (diese lassen sich in einer Klasse, der k-Klasse, zusammenfassen) und in Schätzverfahren, die die Parameter des gesamten Modells simultan schätzen (diese lassen sich ebenfalls in einer Klasse, der K-Matrix-Klasse¹⁾, zusammenfassen). Da die k-Klasse in gewisser Weise ein Spezialfall der K-Matrix-Klasse ist, können alle hier behandelten Schätzverfahren in dieser Oberklasse zusammengefaßt werden.

Zu den in diesem Kapitel auftretenden Begriffen, Schreibweisen usw. vergleiche man Anhang A1.

1. OLS-, GLS- und IV-Methode

Eine Schätzung ist, wie bereits erwähnt, der Schluß von bekanntem Beobachtungsmaterial der ökonomischen Variablen (das im Falle der endogenen Variablen als Stichprobe, im Falle der exogenen Variablen i.a. als fest vorgegeben aufgefaßt wird) auf die unbekannten "wahren" Parameter eines Modells. Gefragt ist nach einem sinnvollen, möglichst optimalen Prinzip oder Verfahren, das die Parameter als Funktion der Beobachtungsmatrizen ausdrückt. Setzt man dann einen bestimmten Beobachtungswertesatz in

1) Scharf, W.: K-Matrix-Class Estimators..., a.a.O.

die Funktion ein, bekommt man einen Schätzwert für den unbekannten Parameter. Die Funktion selbst heißt Schätzfunktion. Ein fundamentales Schätzprinzip ist das Prinzip der kleinsten Quadrate, welches auf Gauß zurückgeht und nicht nur in Zusammenhang mit der Parameterschätzung von Bedeutung ist. Im Zusammenhang mit der Schätzung der Parameter eines interdependenten ökonometrischen Modells wird es im folgenden Unterabschnitt 1.1 als OLS("ordinary least squares")-Methode dargestellt. Bei allgemeineren Voraussetzungen über die Störglieder als in Kapitel I,2. ist die in Unterabschnitt 1.2 dargestellte GLS("generalized least squares")-Methode das geeignete Schätzverfahren. Die IV(Instrumenten-Variablen)-Methode (siehe Unterabschnitt 1.3) ist eine weitere Verallgemeinerung der OLS-Methode, die besonders in Zusammenhang mit interdependenten Systemen von Bedeutung ist. Die OLS-, GLS-, und IV-Methoden lassen sich jedoch nicht nur als Einzelgleichungsschätzverfahren wie im folgenden darstellen, sondern sie sind vielmehr allgemeine Schätzmethoden, mit Hilfe derer man zu komplizierteren Schätzverfahren gelangen kann. So läßt sich z.B. das in Abschnitt 2.2 noch zu behandelnde 2SLS-Schätzverfahren sowohl mit Hilfe der GLS- als auch mit Hilfe der IV-Methode herleiten. Die drei Schätzmethoden OLS, GLS und IV sind daher dem gesamten Kapitel über Schätzverfahren vorangestellt.

1.1. Die OLS-Schätzmethode

Zu schätzen seien die Parameter der i-ten Strukturgleichung (die identifizierbar sei)

$$(1.1) \quad y_i = Y_i \gamma_i + X_i \beta_i + u_i = Z_i \alpha_i + u_i$$

des linearen interdependenten Modells(I,2.1). Die Voraussetzungen und Bezeichnungen sind hier wie im gesamten

Kapitel II die von Kapitel I, 2.. Sie werden im folgenden immer implizit vorausgesetzt, auch wenn sie nicht erwähnt werden.

Die OLS(ordinary least-squares/gewöhnliche Kleinst-Quadrate)-Methode bestimmt einen Schätzwert für α_i so als Funktion der Beobachtungsmatrix (y_i, z_i) , daß die Summe der Quadrate der Störglieder (hier als Residuen betrachtet) minimal wird¹⁾. Es wird also minimiert der Ausdruck

$$u_i' u_i = (y_i - z_i' \alpha_i)' (y_i - z_i' \alpha_i) \text{ als Funktion von } \alpha_i.$$

Es ist $\frac{\partial u_i' u_i}{\partial \alpha_i} = -2 z_i' y_i + 2 z_i' z_i \alpha_i$

Nullsetzen²⁾ und Auflösen nach α_i ergibt

$$(1.2) \quad \hat{\alpha}_i = (z_i' z_i)^{-1} z_i' y_i$$

Definition 1.1: Die Schätzfunktion $\hat{\alpha}_i = (z_i' z_i)^{-1} z_i' y_i$ heißt OLS-Schätzfunktion für α_i aus (1.1).

Eine Aussage über Schätzeigenschaften³⁾ der OLS-Schätzfunktion macht für einen speziellen Fall der folgende Satz:

Satz 1.1: Besteht z_i in Gleichung (1.1) nur aus nichtstochastischen Größen (d.h. $z_i = X_i$, keine verzögerten endogenen Variablen treten in der i-ten Gleichung auf und die exogenen Variablen sind nichtstochastisch), so ist die OLS-Schätzfunktion aus Def. 1.1 ein BLUE (also eine lineare erwartungstreue Schätzfunktion, die innerhalb der Klasse der linearen erwartungstreuen Schätzfunktionen von kleinster Streuung ist).

1) OLS-Schätzung von (1.1) wird auch als multiple Regression bezeichnet. z_i ist die Regressormatrix, y_i der Regressand.

2) Die zweite Bedingung für ein Minimum folgt aus der Konvexität von $u_i' u_i$.

3) vgl. dazu Anhang A1

Beweis: Satz 1.1 folgt als Spezialfall mit $\theta=I$ aus Satz 2.1 im folgenden Abschnitt.

Unter der Voraussetzung, daß Z_i nichtstochastisch ist, besitzt die OLS-Schätzfunktion (1.1) also optimale Schätzeigenschaften. Wie sieht es nun im allgemeinen Fall aus? Da die jetzt in Gleichung (1.1) tatsächlich auftretenden Y_i nicht mehr unkorreliert mit den Störgliedern u_i sind, ist die OLS-Schätzfunktion nicht mehr erwartungstreu. Daß in diesem Fall sogar asymptotische Schätzeigenschaften (d.h. Eigenschaften bei unbegrenzt wachsendem Stichprobenumfang) verloren gehen, zeigt folgender Satz:

Satz 1.2: Falls in der i-ten Strukturgleichung (1.1) erklärende abhängige Variable auftreten, ist die OLS-Schätzfunktion (1.2) inkonsistent.

Beweis: Es ist

$$\hat{\alpha}_i = (Z_i' Z_i)^{-1} Z_i' Y_i = \alpha_i + (Z_i' Z_i)^{-1} Z_i' u_i \quad \text{und} \\ \text{plim } \hat{\alpha}_i = \alpha_i + (\text{plim } T^{-1} Z_i' Z_i)^{-1} (\text{plim } T^{-1} Z_i' u_i),$$

wegen der Eigenschaften des plim-Operators. Nun ist nur noch zu zeigen, daß die beiden plim's auch existieren, wobei - Existenz vorausgesetzt - $\text{plim } T^{-1} Z_i' Z_i$ nichtsingulär sein soll:

$$\text{plim } T^{-1} Z_i' Z_i = \text{plim } T^{-1} \begin{pmatrix} Y_i' Y_i & Y_i' X_i \\ X_i' Y_i & X_i' X_i \end{pmatrix}$$

$\text{plim } T^{-1} Y_i' Y_i = \text{plim } T^{-1} (\Pi_i' X_i' + V_i') (X \Pi_i + V_i)$, wo Π_i und V_i jeweils die Y_i entsprechenden Untermatrizen von Π und V der reduzierten Form sind. Also ist $\text{plim } T^{-1} Y_i' Y_i =$
 $\text{plim } T^{-1} \Pi_i' X_i' X \Pi_i + \text{plim } T^{-1} V_i' X \Pi_i + \text{plim } T^{-1} \Pi_i' X_i' V_i + \text{plim } T^{-1} V_i' V_i$
 $= \Pi_i' M \Pi_i + D_i' \Sigma D_i$, wenn D_i die Y_i entsprechende Untermatrix von Γ^{-1} ist, wegen der Annahmen in Kapitel I über die Störglieder.

$\text{plim } T^{-1} X_1' X_1 = M_{11}$, die entsprechende Untermatrix von M .

$\text{plim } T^{-1} Y_1' X_1 = W_1' M_{11}$, wo M_{11} ebenfalls in naheliegender Weise eine Untermatrix von M ist.

$\text{plim } T^{-1} X_1' Y_1 = M_{11}' W_1$.

Weiter ist

$\text{plim } T^{-1} Z_1' u_1 = \text{plim } T^{-1} \begin{pmatrix} Y_1' u_1 \\ X_1' u_1 \end{pmatrix}$ und

$\text{plim } T^{-1} Y_1' u_1 = \text{plim } T^{-1} W_1' X_1' u_1 + \text{plim } T^{-1} D_1' U' u_1 = D_1' \sigma_1 \neq 0$

$\text{plim } T^{-1} X_1' u_1 = 0$ auf Grund der Annahmen über X .

Alle plim 's existieren also. q.e.d.

Auf Grund dessen, daß die OLS-Schätzfunktion bei einer Gleichung mit erklärenden abhängigen Variablen (also bei tatsächlichem Auftreten von Interdependenz) keine günstigen Schätzeigenschaften besitzt, wurden die in den Abschnitten 2 und 3 zu behandelnden Schätzverfahren entwickelt, die die Interdependenz besser berücksichtigen, und die zumindest asymptotisch bessere Eigenschaften aufweisen. Die Unzulänglichkeit der OLS-Methode für interdependente Systeme wurde als erster von Haavelmo¹⁾ erkannt ("Haavelmo-Bias").

1.2. Die GLS-Schätzmethode

Die GLS (generalized least-squares/verallgemeinerte Kleinst-Quadrate)-Schätzmethode²⁾ ist das genaue Pendant

1) Haavelmo, T.: The Statistical Implications of a System of Simultaneous Equations, *Econometrica* 11, 1943, 1-12.

2) Aitken, A.C.: On Least-Squares and Linear Combination of Observation, in: *Proceedings of the Royal Society of Edinburgh*, vol. 55, 1934, 42-48.

zur OLS-Methode bei einer Verallgemeinerung der stochastischen Annahmen 1(bc)' von Kap. I,2. Beschränkt auf die i-te Strukturgleichung lautet Annahme 1(bc)':

$$1(bc)' \quad E\{u_i u_i'\} = \sigma_{ii} \cdot I$$

Statt I läßt man nun eine beliebige positiv definite Matrix θ zu, d.h. man nimmt an, daß

$$1(bc)^* \quad E\{u_i u_i'\} = \sigma_{ii} \cdot \theta$$

Dies bedeutet, daß nun auch das Störglied einer Gleichung zu verschiedenen Zeitpunkten mit sich selbst korreliert sein darf. Man spricht dann von **A u t o k o r r e l a t i o n**. Das Problem der Autokorrelation wird hier jedoch nicht weiter erörtert, und die GLS-Schätzfunktion wird unter den verallgemeinerten stochastischen Voraussetzungen nur hergeleitet in Hinblick auf ihre Bedeutung für die Entwicklung der später noch zu behandelnden Schätzfunktionen, für die natürlich wieder die ursprünglichen stochastischen Annahmen aus Kap. I,2 gelten.

Die Frage ist nun die, ob sich unter den verallgemeinerten stochastischen Annahmen 1(bc)* eine ähnlich gute Schätzfunktion für α_i aus (1.1) finden läßt, wie es im Fall der Nicht-Autokorrelation bei nichtstochastischem Z_i die OLS-Schätzfunktion ist. Um zur OLS-Schätzfunktion zu gelangen, hatte man in Abschnitt 1.1 den Ausdruck $u_i' u_i$ oder gleichbedeutend damit $u_i' I^{-1} u_i$ minimiert als Funktion von α_i . In der verallgemeinerten Annahme 1(bc)* wurde aber gerade die Matrix I der Annahme 1(bc)' durch die positiv definite Matrix θ ersetzt. Es liegt also nahe, auch im Minimierungsprozeß I durch θ zu ersetzen und den Ausdruck

$$u_i' \theta^{-1} u_i$$

als Funktion von α_i zu minimieren.

Differenzieren und Nullsetzen ergibt folgende Schätzfunktion:

$$(1.3) \quad \hat{\alpha}_1 = (Z_1' \theta^{-1} Z_1)^{-1} Z_1' \theta^{-1} y_1$$

Definition 1.2: Die Schätzfunktion $\hat{\alpha}_1 = (Z_1' \theta^{-1} Z_1)^{-1} Z_1' \theta^{-1} y_1$ heißt GLS(oder Aitken-) - Schätzfunktion für α_1 aus (1.1) unter den verallgemeinerten stochastischen Voraussetzungen 1(bc)*.

Setzt man $\theta = I$, ergibt sich die OLS-Schätzfunktion als Spezialfall der GLS-Schätzfunktion. Der Beweis, daß die OLS-Schätzfunktion ein BLUE ist, wird daher erst hier als Spezialfall des folgenden allgemeineren Beweises ausgeführt.

Satz 1.3: Besteht Z_1 in Gleichung (1.1) nur aus nichtstochastischen Größen, so ist die GLS-Schätzfunktion aus Definition 1.2 ein BLUE unter der verallgemeinerten stochastischen Annahme 1(bc)*.

Beweis: Sei $\tilde{\alpha}_1 = C'y_1$ irgendeine lineare Schätzfunktion von α_1

$$\Rightarrow \tilde{\alpha}_1 = C'Z_1\alpha_1 + C'u_1 \quad \text{und}$$

$$E\{\tilde{\alpha}_1\} = C'Z_1\alpha_1$$

Soll nun $\tilde{\alpha}_1$ erwartungstreu sein, so muß folgende Bedingung erfüllt sein:

$$(1.4) \quad C'Z_1 = I$$

Weiter ist die Kovarianzmatrix der geschätzten α_1 unter dieser Bedingung (1.4) gleich

$$\Sigma_{\tilde{\alpha}_1, \tilde{\alpha}_1} = E\{(\tilde{\alpha}_1 - \alpha_1)(\tilde{\alpha}_1 - \alpha_1)'\} = E\{C'u_1 u_1' C'\} , \text{ da}$$

$$\tilde{\alpha}_1 - \alpha_1 = C'u_1 \quad \text{und} \quad E\{C'u_1 u_1' C'\} = \sigma_{11} C' \theta C .$$

Man betrachte nun die GLS-Schätzfunktion aus Definition 1.2:

Sie ist linear mit $\hat{C}' = (Z_1' \theta^{-1} Z_1)^{-1} Z_1' \theta^{-1}$

Sie ist erwartungstreu wegen $C'Z_1 = I$ (s. Bed.(1.4))

Ihre Kovarianzmatrix ist

$$\Sigma_{\hat{\alpha}_1 \hat{\alpha}_1} = \sigma_{11} \hat{C}' \theta \hat{C} = \sigma_{11} (Z_1' \theta^{-1} Z_1)^{-1}$$

Sei nun C' von irgendeiner linearen erwartungstreuen Schätzfunktion. Dann kann man C' ausdrücken als

$$C' = \hat{C}' + D' \quad \text{mit} \quad D'Z_1^* = 0 \quad \text{wegen Bedingung (1.4)}$$

$$\begin{aligned} \Rightarrow \Sigma_{\tilde{\alpha}_1 \tilde{\alpha}_1} &= E\{C' u_1 u_1' C\} \\ &= E\{(\hat{C}' + D') u_1 u_1' (\hat{C} + D)\} \\ &= \sigma_{11} (\hat{C}' + D') \theta (\hat{C} + D) \\ &= \sigma_{11} (\hat{C}' \theta \hat{C} + D' \theta D + \hat{C}' \theta D + D' \theta \hat{C}) \\ &= \sigma_{11} (\hat{C}' \theta \hat{C} + D' \theta D) \quad , \text{ da } D'Z_1 = 0 \\ &= \sigma_{11} \Sigma_{\hat{\alpha}_1 \hat{\alpha}_1} + \sigma_{11} D' \theta D \quad \text{mit positiv definitem } D' \theta D. \end{aligned}$$

Also ist $\hat{\alpha}_1$ tatsächlich von kleinster Kovarianzmatrix q.e.d.

Im allgemeinen Fall, also bei stochastischem Z_1 in Gleichung (1.1) gilt Satz 1.3 natürlich nicht mehr. Darauf soll hier jedoch nicht näher eingegangen werden, sondern im nächsten Abschnitt eine weitere Schätzmethode dargestellt werden, die für die Konstruktion der Schätzverfahren der Abschnitte 2. und 3. von Bedeutung sein wird.

1.3. Die IV-Schätzmethode

Das IV (instrumental variable/Instrumenten-Variablen)-Schätzverfahren ist eine allgemeine Schätzmethode, die eine konsistente Schätzfunktion für α_1 aus (1.1) quasi erzwingt. Aus dem Beweis der Inkonsistenz der OLS-Schätz-

funktion bei stochastischem Z_i sieht man, daß der Grund für die Inkonsistenz der war, daß $\text{plim } T^{-1}Z_i'u_i \neq 0$ war, also auch asymptotisch noch ein stochastischer Zusammenhang zwischen den erklärenden Variablen und dem Störglied bestand. Beim IV-Schätzverfahren nimmt man nun an, daß man über einen Beobachtungswertesatz von $n_i + r_i$ Hilfsvariablen (Instrumentenvariablen) verfügt, der in der $Tx(n_i + r_i)$ -Matrix P_i zusammengefaßt ist, und der so beschaffen ist, daß einerseits

$$(1.5) \quad \text{plim } T^{-1}P_i'u_i = 0,$$

andererseits jedoch ein stochastischer Zusammenhang zwischen P_i und Z_i besteht, insoweit als

$$(1.6) \quad \text{plim } T^{-1}P_i'Z_i = M_i \quad \text{als nichtsinguläre Matrix}$$

existiert.

Der Beobachtungswertesatz P_i muß aus Beobachtungswerten aus dem Modell bestehen. Wie man ihn konkret konstruiert, wird später noch an einzelnen Fällen gezeigt. Die IV-Schätzfunktion wird nun in Anlehnung an die OLS-Schätzfunktion definiert als

$$(1.7) \quad \hat{\alpha}_i = (P_i'Z_i)^{-1}P_i'y_i.$$

Definition 1.3: Die Schätzfunktion $\hat{\alpha}_i = (P_i'Z_i)^{-1}P_i'y_i$ mit einer Hilfsvariablenmatrix P_i , die die Bedingungen

$$\begin{aligned} \text{plim } T^{-1}P_i'u_i &= 0 \quad \text{und} \\ \text{plim } T^{-1}P_i'Z_i &= M_i \quad (|M_i| \neq 0) \end{aligned}$$

erfüllt, heißt IV-Schätzfunktion für α_i aus (1.1).

Unmittelbar ergibt sich aus dem oben Gesagten der fol-

gende Satz:

Satz 1.4: Die IV-Schätzfunktion (1.7) aus Definition 1.3 ist eine konsistente Schätzfunktion für α_1 .

Beweis: $\text{plim } \hat{\alpha}_1 = \alpha_1 + (\text{plim } T^{-1}P_1'Z_1)^{-1}\text{plim } T^{-1}P_1'u_1$
 $= \alpha_1$, da $\text{plim } T^{-1}P_1'Z_1$ nach Annahme (1.6)

existiert und nichtsingulär ist, und da $\text{plim } T^{-1}P_1'u_1 = 0$ nach (1.5), q.e.d.

Ein (triviales) Beispiel für die konkrete Wahl der Hilfsvariablen in einem speziellen Fall ist das folgende: In Gleichung (1.1) sei $Z_1 = X_1$. Als Hilfsvariablenmatrix wähle man X_1 selbst. Wegen der Annahmen 2a) und b) in I,2 erfüllt X_1 in der Tat die Bedingungen (1.5) und (1.6). Als IV-Schätzfunktion für α_1 ergibt sich dann

$\hat{\alpha}_1 = (X_1'X_1)^{-1}X_1'y_1$, die OLS-Schätzfunktion.

Zu einer weiteren ausführlichen Diskussion der IV-Schätzfunktion siehe Sargan¹⁾.

2. Die Schätzverfahren der k-Klasse

Nachdem in Abschnitt 1 einige grundsätzliche Schätzprinzipien erörtert wurden, werden nun die Schätzverfahren behandelt, die speziell für interdependente Modelle unter besonderer Berücksichtigung der bei Interdependenz auftretenden Besonderheiten entwickelt wurden. Die Schätz-

1) Sargan, J.D.: The Estimation of Economic Relationships Using Instrumental Variables, *Econometrica* 26, 1958, 393-415.

verfahren, die speziell zur Schätzung der Parameter einer Strukturgleichung eines interdependenten Modells entwickelt wurden, lassen sich in der sogenannten k-Klasse zusammenfassen, die durch Variation eines freien Parameters (nämlich k) erzeugt wird. Man kann also von den folgenden Schätzverfahren auch von Spezialfällen (bei einer festen Wahl von k) einer einparametrischen Schar von Schätzverfahren sprechen. Um sie besser Motivieren zu können, erfolgt die genaue Definition der k-Klasse erst in Unterabschnitt 2.4, wenn ihre Spezialfälle erörtert sind. Da bei den Einzelgleichungsschätzverfahren der k-Klasse bei der Schätzung nicht die volle Information über das Modell ausgeschöpft wird, spricht man auch von *l i m i t e d - i n - f o r m a t i o n* -Schätzverfahren (Schätzverfahren bei begrenzter Information).

2.1. Die ILS-Methode

Das ILS (indirect least-squares/indirekte Kleinst-Quadrate)- Schätzverfahren versucht (naheliegenderweise) zur Schätzung der Strukturparameter die bei Identifizierbarkeit einer Gleichung bestehende eindeutige Beziehung zwischen reduzierter Form und Strukturform (vgl. Kap. I, Satz 3.3) auszunutzen und von den geschätzten reduzierten-Form-Parameter zu einer Schätzung der Strukturparameter zu gelangen. Daß dies jedoch nur in einem sehr speziellen Fall möglich ist, zeigt das Folgende.

Zu schätzen seien wieder die Parameter der i-ten Strukturgleichung (1.1), die hier nochmals aufgeschrieben sei:

$$(2.1) \quad y_i = Y_i \gamma_i + X_i \beta_i + u_i = Z_i \alpha_i + u_i$$

Die i-te Gleichung der reduzierten Form lautet (vgl. I, (3.2))

$$y_i = X \pi_i + v_i$$

Aus der Diskussion in 1.1 erkennt man, daß die OLS-Schätzfunktion von π_1 bei nichtstochastischem X ein BLUE ist. Beim Auftreten von verzögerten endogenen Variablen in X ist die OLS-Schätzfunktion von π_1 zumindest noch konsistent (wegen I, 2, Ann. 2(a),(b)). Schätzt man die Parameter aller N reduzierten-Form-Gleichungen mit der OLS-Methode und faßt die geschätzten Parametervektoren $\hat{\pi}_1$ zur Matrix $\hat{\Pi}$ zusammen, so bekommt man eine Schätzung für Π , die entsprechend den Annahmen über X ein BLUE oder konsistent ist.

Nach I, (3.3') besteht bei Identifizierbarkeit folgende (bis auf Normierung) eineindeutige Beziehung zwischen reduzierte-Form-Parametern und den Parametern der i -ten Strukturgleichung:

$$(2.2) \quad \begin{pmatrix} \pi_{11} & I_1 \\ \pi_{21} & 0 \end{pmatrix} \begin{pmatrix} \bar{\delta}_1 \\ \beta_1 \end{pmatrix} = 0$$

Entsprechend Gleichung (2.1) wird das obige Gleichungssystem (2.2) durch $\bar{\delta}_{11} = -1$ normiert. Setzt man nun für π_{11} und π_{21} ihre OLS-Schätzwerte $\hat{\pi}_{11}$ und $\hat{\pi}_{21}$ ein und löst

das Gleichungssystem nach $\begin{pmatrix} \bar{\delta}_1 \\ \beta_1 \end{pmatrix}$ auf, so bekommt man ein-

deutige Schätzwerte für γ_1 und β_1 , wobei allerdings folgendes zu beachten ist: Da π_{21} aus Schätzwerten besteht, die ohne Berücksichtigung der a-priori Restriktionen der Strukturparameter (nämlich durch OLS-Schätzung der reduzierten Form) gewonnen wurden, wird sich die Rangbedingung (vgl. I, Satz 3.2) für die eindeutige Auflösbarkeit von (2.2) nur dann erfüllen lassen, wenn π_{21} maximalen Rang hat, wenn also $\text{rg}(\pi_{21}) = N_1 - 1 = \bar{R}_1$ gilt. Wenn man in (2.2) π_{11} und π_{21} durch $\hat{\pi}_{11}$ und $\hat{\pi}_{21}$ ersetzt, wird sich das Gleichungssystem also nur dann eindeutig auflösen lassen (und damit eindeutige Schätzwerte für γ_1 und β_1

erzeugen), wenn $N_i - 1 = \bar{R}_i$ ist, also der Fall der genauen Identifizierbarkeit der i-ten Strukturgleichung vorliegt.

Daraus ergibt sich folgende Definition der ILS-Schätzfunktion:

Definition 2.1: Die i-te Strukturgleichung (2.1) des interdependenten Modells sei genau identifizierbar. Dann heißt die Schätzfunktion $\hat{\alpha}_i$, die sich durch Auflösung des Gleichungssystems

$$\begin{pmatrix} \hat{\Pi}_{11} & I_1 \\ \hat{\Pi}_{21} & 0 \end{pmatrix} \begin{pmatrix} \bar{\delta}_i \\ \beta_i \end{pmatrix} = 0 \quad \text{unter der Normierung } \bar{\delta}_{ii} = -1$$

ergibt, ILS-Schätzfunktion für α_i . $\hat{\Pi}_{11}$ und $\hat{\Pi}_{21}$ sind dabei Untermatrizen der OLS-geschätzten reduzierten Form $\hat{\Pi}$.

Über die Schätzeigenschaften läßt sich folgendes sagen:

Satz 2.1: Die ILS-Schätzfunktion $\hat{\alpha}_i$ aus Definition 2.1 ist konsistent.

Beweis: Folgt unmittelbar aus der Konsistenz von $\hat{\Pi}$.

2.2. Die 2SLS-Methode

Wie 2.1. gezeigt hat, ist die ILS-Methode ein Schätzverfahren, das zwar die Interdependenz des Modells besser als OLS berücksichtigt, das jedoch nur im Fall der genauen Identifizierbarkeit anwendbar ist. In der Praxis wird jedoch typischerweise der Fall der Überidentifizierbarkeit auftreten. Ein Schätzverfahren, das auch in diesem allgemeinen Fall anwendbar ist, ist das 2SLS (two-stage least-squares/zweistufige Kleinst-Quadrate)-Schätzverfahren, das von Theil¹⁾ und Basman²⁾ entwickelt wurde. Es

1) Theil, H.: Economic Forecasts and Policy, 1st ed., Amsterdam 1958.

2) Basman, R.L.: A Generalized Classical Method of Linear Estimation of Coefficients in a Structural Equation, Econometrica 25, 1957, 77-83.

läßt sich auf verschiedene Weisen herleiten und motivieren.

2.2.1. 2SLS als IV-Schätzung

Zu schätzen ist wieder die i-te Strukturgleichung (2.1). Es sollen Hilfsvariablen P_i gefunden werden, so daß

$$\begin{aligned} \text{plim } T^{-1}P_i'u_i &= 0 \\ \text{plim } T^{-1}P_i'Z_i &= M_i \quad (|M_i| \neq 0). \end{aligned}$$

Die daraus resultierende IV-Schätzfunktion ist nach Satz 1.4 konsistent. Als Instrumente P_i für Z_i wähle man

$$P_i = (\hat{Y}_i, X_i) = (X\hat{\Pi}_i, X_i) = (X(X'X)^{-1}X'Y_i, X_i) = X(X'X)^{-1}X'Z_i$$

aus der Überlegung heraus, daß zwischen $X\Pi_i$ und u_i ein genügend schwacher stochastischer Zusammenhang besteht und die OLS-Schätzung $\hat{\Pi}_i$ für Π_i (die Y_i entsprechende Untermatrix von Π) eine konsistente Schätzung ist. Daß die Bedingungen $\text{plim } T^{-1}P_i'u_i = 0$ und $\text{plim } T^{-1}P_i'Z_i = M_i$ ($|M_i| \neq 0$) mit dieser Wahl der Instrumente tatsächlich erfüllt sind, läßt sich auf Grund der Annahmen I, 2(a), (b) leicht zeigen. Auf einen Beweis wird daher hier verzichtet, zumal er später in einem allgemeineren Zusammenhang für k-Klasse-Schätzungen noch erbracht wird.

Betrachtet man die IV-Schätzfunktion (1.6) und setzt in sie unsere Wahl von P_i ein, so erhält man:

$$(2.3) \quad \hat{\alpha}_i = (Z_i'X(X'X)^{-1}X'Z_i)^{-1}Z_i'X(X'X)^{-1}X'Y_i$$

und es ergibt sich folgende Definition:

Definition 2.2: Die Schätzfunktion

$$\hat{\alpha}_i = (Z_i'X(X'X)^{-1}X'Z_i)^{-1}Z_i'X(X'X)^{-1}X'y_i$$

heißt 2SLS-Schätzfunktion für α_i aus (2.1).

Die wichtigste Schätzeigenschaft für die 2SLS-Schätzfunktion ergibt sich unmittelbar aus ihrer Interpretation als IV-Schätzung:

Satz 2.2: Die 2SLS-Schätzfunktion $\hat{\alpha}_i$ aus Definition 2.2 ist konsistent.

Es läßt sich sogar zeigen, daß 2SLS eine "optimale" IV-Schätzfunktion ist¹⁾.

Bemerkung: Von einem "zweistufigen" Schätzverfahren spricht man, weil man in zwei Schätzschritten vorgeht: Zuerst die OLS-Schätzung von Π_i , dann die IV-Schätzung mit der OLS-Schätzung von Π_i als Instrument. Wegen der Idempotenz²⁾

von $X(X'X)^{-1}X'$ gilt $P_i'Z_i = P_i'P_i$. Die IV-Schätzung läßt sich also sogar als OLS-Schätzung des Ansatzes $y_i = P_i'\alpha_i + u_i^*$ betrachten. 2SLS resultiert dann tatsächlich aus zweimaliger Anwendung von OLS.

2.2.2. 2SLS als GLS-Schätzung

Eine zweite Interpretation des 2SLS-Schätzverfahrens ergibt sich aus folgender Überlegung:

Der Grund dafür, daß man Gleichung (2.1) nicht konsistent OLS schätzen kann, liegt ja daran, wie der Beweis von Satz 1.2 zeigt, daß ein stochastischer Zusammenhang zwi-

1) vgl. Dhrymes, P.J.: Econometrics, a.a.O. S.302.

2) Eine Matrix A heißt idempotent, wenn $AA=A$.

schen Y_i und u_i auch asymptotisch noch besteht ($\text{plim} T^{-1} Y_i' u_i \neq 0$). Nun könnte man versuchen, diesen asymptotischen stochastischen Zusammenhang dadurch auszuschließen, daß man die gesamte Gleichung (2.1) geeignet transformiert. Nimmt man als Transformationsmatrix die Matrix der Vorherbestimmten X , so erhält man

$$(2.4) \quad X'Y_i = X'Y_i\delta_i + X'X_i\beta_i + X'u_i = X'Z_i\alpha_i + X'u_i \quad .$$

Da man weiß, daß $\text{plim} T^{-1} X'Y_i$ existiert (vgl. Beweis von Satz 1.2) - $T^{-1} X'Y_i$ also der Wahrscheinlichkeit nach gegen einen nichtstochastischen Wert konvergiert -, kann man zumindest asymptotisch erwarten, daß kein stochastischer Zusammenhang mehr zwischen $X'Y_i$ und $X'u_i$ besteht und damit die Kleinst-Quadrate Schätzung der transformierten Gleichung (2.4) konsistent ist. Allerdings kann man dabei nicht auf OLS zurückgreifen, da nun für das transformierte Störglied $X'u_i$ nicht mehr die Bedingung I, Annahme 1(c) gilt. Man verwendet vielmehr (zumindest formal) GLS, wobei man so tut, als sei X nichtstochastisch und damit

$$E\{X'u_i u_i' X'\} = X'\sigma_{ii}IX = \sigma_{ii}X'X \quad .$$

Man hat also den Fall von Annahme 1(bc)* aus Abschnitt 1.2 vorliegen mit $\theta = X'X$, und man kann nun mit diesem θ die Schätzformel (1.3) auf die transformierte Gleichung (2.4) anwenden und erhält (das Z_i aus (1.3) ist jetzt $X'Z_i$)

$$\hat{\alpha}_i = (Z_i'X(X'X)^{-1}X'Z_i)^{-1}Z_i'X(X'X)^{-1}X'y_i \quad ,$$

was genau der 2SLS-Schätzfunktion (2.3) für α_i entspricht.

Obige Diskussion zeigt also, daß man (formal) die 2SLS-Schätzfunktion für die Parameter der i -ten Strukturgleichung als GLS-Schätzfunktion der mit X transformierten

Strukturgleichung betrachten kann. "Formal" heißt dabei, daß man bei der Herleitung der Kovarianzmatrix des transformierten Störgliedes $X'u_i$ so tut, als sei X nichtstochastisch. Eine echte GLS-Schätzung liegt also nur im Fall von nichtstochastischem X vor. Satz 2.2 zeigt jedoch, daß die daraus resultierende Schätzung (nämlich 2SLS) in beiden Fällen konsistent ist.

2.2.3. 2SLS als Ergebnis von Minimierungsprozessen

Wie Abschnitt 1.2. gezeigt hat, kann man die GLS-Schätzfunktion als Ergebnis eines Minimierungsprozesses betrachten, nämlich der Minimierung von $u_i'\theta^{-1}u_i$ als Funktion der unbekannten Parameter. Nun ist aber nach 2.2.2. auch die 2SLS-Schätzfunktion eine GLS-Schätzung eines transformierten Ansatzes der Strukturgleichung mit (vgl. (2.4)) $u_i^* = X'u_i$ und $\theta = X'X$. Die 2SLS-Schätzfunktion für die Parameter der i -ten Strukturgleichung (2.1) resultiert also aus der Minimierung von

$$(2.5) \quad u_i'X(X'X)^{-1}X'u_i \quad \text{als Funktion der unbekannten Parameter.}$$

Ersetzt man u_i entsprechend (2.1) durch die unbekannten Parameter α_i , so minimiert man den Ausdruck

$$(2.6) \quad (y_i - Z_i\alpha_i)'X(X'X)^{-1}X'(y_i - Z_i\alpha_i)$$

oder gleichbedeutend damit

$$(2.7) \quad (y_i - \tilde{Z}_i\alpha_i)'(y_i - \tilde{Z}_i\alpha_i) \quad \text{mit} \quad \tilde{Z}_i = X(X'X)^{-1}X'Z_i$$

(Dies folgt durch Ausmultiplizieren von (2.6) und daraus, daß $y_i'X(X'X)^{-1}X'y_i$ nicht von α_i abhängt und $X(X'X)^{-1}X'$ idempotent ist.)

Der Minimand (2.7) ist ein Spezialfall der von

Dhrymes¹⁾ angegebenen Determinante, aus deren Minimierung die später noch zu behandelnde 3SLS-Schätzfunktion resultiert.

Minimierung von (2.5) ist gleichbedeutend mit der Minimierung von

$$(2.8) \quad \text{tr} (U'X(X'X)^{-1}X'U)$$

als Funktion der unbekannten Parameter der i-ten Strukturgleichung und unter Beachtung der a-priori Nullrestriktionen. Es ist nämlich

$$\text{tr} (U'X(X'X)^{-1}X'U) = \sum_{k=1}^N u_k'(\alpha_k) X(X'X)^{-1} X' u_k(\alpha_k)$$

und der Differentialquotient nach α_i ist nur für den Summanden $k=i$ ungleich Null.

Einsetzen von $U=ZA$ mit $Z=(Y,X)$, $A=\begin{pmatrix} \Gamma \\ B \end{pmatrix}$ in (2.8) ergibt:

$$(2.9) \quad \text{tr} (A'Z'ZA) \quad , \quad \text{wo } \tilde{Z} = X(X'X)^{-1}X'Z \quad .$$

(2.9) ist der von Dhrymes¹⁾ angegebene Minimand, wenn die Spur einer Matrix minimiert werden soll. 2SLS erhält man also entweder durch Minimierung (des Spezialfalles) einer Determinante ((2.5) bzw. (2.7)) oder durch Minimierung der Spur einer Matrix ((2.8) bzw. (2.9))²⁾.

1) vgl. Dhrymes, P.J.: Small Sample and Asymptotic Relations between Maximum Likelihood and Three Stage Least Squares Estimators, *Econometrica* 41, 1973, 357-364.

2) vgl. im Zusammenhang mit Minimierungsprozessen die Minimum-Distance-Estimator von Malinvaud: Malinvaud, E.: *Statistical Methods of Econometrics*, a.a.O. S.277, 586ff. Eine geometrische Deutung der Minimierungsprozesse gibt auch Chow, G.C.: A Comparison of Alternative Estimators for Simultaneous Equations, *Econometrica* 32, 1964, 532-553.

2.3. LISE-Methoden

Eine zweite Gruppe von Einzelgleichungsschätzverfahren bei begrenzter Information sind die LISE (limited-information single-equation)- Schätzmethoden, die mit Hilfe von verschiedenen Schätzprinzipien zu denselben Schätzgleichungen und damit zur gleichen Schätzfunktion für die unbekannten Parameter der i-ten Strukturgleichung gelangen. Historisch gesehen ist das auf grund des Maximum-Likelihood-Prinzips entwickelte LISE-Schätzverfahren (LIML) älter¹⁾ als das 2SLS-Verfahren. Wegen der rechnerisch jedoch recht schwierigen Herleitung und Anwendung der Schätzgleichungen wird in der Praxis heute meist das 2SLS-Verfahren vorgezogen. Auch die LISE-Schätzungen gehören zur k-Klasse.

2.3.1. Die LIML-Methode

Dem Maximum-Likelihood-Prinzip liegt folgender Gedanke zu grunde: Zur Bestimmung der unbekannten Parameter des gesamten Modells oder einer Gleichung wird die sogenannte Likelihood-Funktion maximiert. Die Likelihood-Funktion ist formal nichts anderes als die gemeinsame Dichtefunktion der $y(t)$ (der spaltenweise geschriebenen Reihen von Y), aufgefaßt jedoch als Funktion der Parameter, wobei die Störglieder als normalverteilt vorausgesetzt werden: $u(t) \sim N(0, \Sigma)$. Maximieren der Likelihood-Funktion bedeutet

1) vgl. Anderson, T.W. und H. Rubin: Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations, Annals of Mathematical Statistics 20, 1949, 46-63.

Anderson, T.W.: Estimation of the Parameters of a Single Equation by the Limited-Information Maximum-Likelihood Method, Cowles Commission Monograph 10, 1950, S. 311-322.

also, Schätzwerte für die Parameter so zu bestimmen, daß die Wahrscheinlichkeit für das Auftreten der durch die Beobachtungswerte vorgegebenen Stichprobe maximal wird .

Aus der Annahme der Normalverteilung der Störglieder und unter den Annahmen I,1 über die Störglieder erhält man folgende gemeinsame (bedingte) Dichtefunktion der $y(t)$:

$$(2.10) \quad f(y/A, \Sigma) = c_1 \|\Gamma\|^T |\Sigma|^{-T/2} \exp\left\{-\frac{1}{2} \text{tr}\left(\frac{1}{T} Z A \Sigma^{-1} A' Z'\right)\right\}$$

wobei A und Z wie in I, (2.1') erklärt sind. c_1 ist eine Konstante, die hier nicht weiter interessiert. $\|\Gamma\|^T$ ist die Jacobische Determinante, die sich bei der Transformation der Dichtefunktion der $u(t)$ in die Dichtefunktion der $y(t)$ ergibt. Durch logarithmieren²⁾ von (2.10) und dividieren durch T erhält man aus (2.10) folgende log-Likelihood-Funktion:

$$(2.11) \quad L(A, \Sigma) = c_2 + \log \|\Gamma\| - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \text{tr}\left(\frac{1}{T} Z A \Sigma^{-1} A' Z'\right)$$

(2.11) bildet die Ausgangslikelihoodfunktion sowohl für das später noch zu behandelnde FIML-Schätzverfahren, als auch für das in diesem Abschnitt zu entwickelnde LIML (limited-information maximum-likelihood)-Schätzverfahren. Beim LIML-Schätzverfahren werden, wie bei allen Einzelgleichungsschätzverfahren, nur die a-priori Informationen über die zu schätzende Strukturgleichung berücksichtigt. Ohne Beschränkung der Allgemeinheit sei die erste Gleichung zu schätzen. Entsprechend lassen sich die Parameter A und Σ aufspalten in

$$A = (\tilde{a}_1, A_1), \quad \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_1' \\ \sigma_1 & \Sigma_1 \end{pmatrix} .$$

2) Da $\log$ eine monoton wachsende Funktion ist, nehmen Likelihood-Funktion und logarithmierte Likelihood-Funktion an derselben Stelle, also für dieselben Parameterwerte ihr Maximum an.

Das Ziel ist nun, von (2.11) zu einer "konzentrierten" Likelihood-Funktion zu gelangen, die nur noch die Parameter der ersten Gleichung enthält. Diese konzentrierte Likelihood-Funktion ist dann unter Berücksichtigung der a-priori Informationen über die erste Gleichung zu maximieren.

Eine konzentrierte Likelihood-Funktion erhält man, wenn man das gesamte Modell

$$ZA + U = 0$$

transformiert, und zwar so, daß einerseits bei der Transformation der Parametervektor $\bar{\alpha}_1$ der ersten Gleichung erhalten bleibt, andererseits die erste Gleichung durch die Transformation stochastisch unabhängig von dem Restsystem wird, daß also der $(N-1) \times 1$ -Spaltenvektor $\sigma_1 = 0$ wird. Die gesuchte Transformationsmatrix G soll also die Bedingungen

$$(2.12) \quad AG = (\bar{\alpha}_1, A_1^*) \quad \text{sowie} \quad G' \Sigma G = \begin{pmatrix} \sigma_{11} & 0 \\ 0 & \Sigma_1^* \end{pmatrix}$$

erfüllen. Die Matrix

$$(2.13) \quad G = \begin{pmatrix} 1 & , & -\sigma_{11}^{-1} \sigma_1' G_1 \\ 0 & , & G_1 \end{pmatrix} \quad ,$$

wobei G_1 so gewählt ist, daß es

$$G_1' (\Sigma_1 - \sigma_{11}^{-1} \sigma_1 \sigma_1') G_1 = I$$

erfüllt, ist die gesuchte Matrix. Zunächst läßt sich in der Tat ein G_1 finden, das obige Gleichung erfüllt: Mit Σ ist nämlich auch $\Sigma_1 - \sigma_{11}^{-1} \sigma_1 \sigma_1'$ positiv definit, läßt sich also darstellen in der Form $G_1^{-1} G_1^{-1}$. Nun zeigt man leicht

durch einfaches Ausmultiplizieren, daß das G aus (2.13) die Bedingung (2.12) erfüllt, sogar mit $\Sigma_1 = I$.

Durch einige einfache Rechnungen erhält man folgenden Ausdruck für die Log-Likelihood-Funktion des mit G aus (2.13) transformierten Systems:

$$(2.14) \quad L(A, G, \Sigma) = c_2 + \log \|FG\| - \frac{1}{2} \log |G' \Sigma G| \\ - \frac{1}{2} \text{tr} \left(\frac{1}{T} (ZAG) (G' \Sigma G)^{-1} (ZAG)' \right)$$

d.h. $L(A, G, \Sigma)$ ist einfach $L(AG, G' \Sigma G)$ mit dem L aus (2.11).

Einsetzen von G aus (2.13) in (2.14) ergibt

$$L(\dots) = c_2 + \log \|(\bar{\gamma}_1, \Gamma_1)\| - \frac{1}{2} \log \sigma_{11} \\ - \frac{1}{2} \text{tr} \left(\begin{pmatrix} \sigma_{11}^{-1} & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} \bar{\alpha}_1' \\ A_1^* \end{pmatrix} M (\bar{\alpha}_1, A_1^*) \right) \quad , \text{ wo}$$

$$AG = \begin{pmatrix} \bar{\alpha}_1 & \Gamma_1 \\ \beta_1 & B_1 \end{pmatrix} = (\bar{\alpha}_1, A_1^*) \quad \text{und} \quad M = \frac{1}{T} Z' Z \quad . \quad M \text{ ist natürlich}$$

nicht mit dem M aus I, Annahme 2(a) zu verwechseln. Der Querstrich über $\bar{\gamma}_1$ zeigt, daß noch keine Restriktionen auferlegt wurden (insbesondere auch keine Normierung). Weitere Umformung ergibt:

$$(2.15) \quad L(\dots) = c_2 + \log \|(\bar{\gamma}_1, \Gamma_1)\| - \frac{1}{2} \log \sigma_{11} \\ - \frac{1}{2} \sigma_{11}^{-1} \bar{\alpha}_1' M \bar{\alpha}_1 - \frac{1}{2} \text{tr} (A_1^{*'} M A_1^*) \quad .$$

Nun sind die Parameter der interessierenden ersten Gleichung schon fast von den übrigen getrennt, mit Ausnahme des Ausdrucks $\log \|(\bar{\gamma}_1, \Gamma_1)\|$.

Um A_1^* aus (2.15) zu eliminieren, maximiert man L in Bezug auf A_1^* , d.h. man differenziert (2.15) partiell nach

A_1^* und setzt den so erhaltenen Ausdruck gleich Null¹⁾.
Es ist

$$\frac{\partial}{\partial A_1^*} \log \|(\bar{y}_1, \Gamma_1) \| = \left(\frac{\frac{\partial \log \|(\bar{y}_1, \Gamma_1) \|}{\partial \Gamma_1}}{\frac{\partial \log \|(\bar{y}_1, \Gamma_1) \|}{\partial B_1}} \right) = \begin{pmatrix} H_1 \\ 0 \end{pmatrix}, \text{ wo } H_1$$

definiert ist durch $H = (h_1, H_1) = \begin{pmatrix} \bar{y}_1^1 \\ \Gamma_1^1 \end{pmatrix}^{-1}$.

Weiter ist

$$\frac{d}{dA_1^*} \left\{ \frac{1}{2} \text{tr}(A_1^* M A_1^*) \right\} = M A_1^*.$$

Die anderen Terme in (2.15) hängen nicht von A_1^* ab.

Nullsetzen von $\partial L / \partial A_1^*$ ergibt also

$$(2.16) \quad \begin{pmatrix} H_1 \\ 0 \end{pmatrix} - M A_1^* = 0.$$

Daraus erhält man

$$(2.17) \quad H_1 = W \Gamma_1 \quad \text{mit} \quad W = \frac{1}{T} Y' (I - X(X'X)^{-1} X') Y,$$

wie man durch einsetzen von $W \Gamma_1$ in (2.16) und eliminieren von B_1 leicht zeigt. Wegen (2.16) hat man

$$\begin{aligned} \text{tr}(A_1^* M A_1^*) &= \text{tr} \left(\begin{pmatrix} \Gamma_1^1 & B_1^1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} H_1 \\ 0 \end{pmatrix} \right) = \text{tr}(\Gamma_1^1 H_1) = \\ (2.18) \quad &= \text{tr} \left((0, I) \begin{pmatrix} \bar{y}_1^1 \\ \Gamma_1^1 \end{pmatrix} \begin{pmatrix} \bar{y}_1^1 \\ \Gamma_1^1 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ I \end{pmatrix} \right) = \text{tr}(I) = N-1 \end{aligned}$$

Weiter ist wegen (2.17)

$$(\bar{y}_1, \Gamma_1)' W (\bar{y}_1, \Gamma_1) = \begin{pmatrix} \bar{y}_1^1 W \bar{y}_1 & \bar{y}_1^1 W \Gamma_1 \\ \Gamma_1^1 W \bar{y}_1 & \Gamma_1^1 W \Gamma_1 \end{pmatrix}$$

mit

1) Hier wie bei den anderen Maximierungs- bzw. Minimierungsprozessen folgt die 2. Bedingung für einen Extremwert jeweils aus der Konvexität bzw. Konkavität der Funktion.

$$\text{mit } \bar{\gamma}_1' W \Gamma_1 = \bar{\gamma}_1' H_1 = 0$$

$$\Gamma_1' W \Gamma_1 = \Gamma_1' H_1 = I$$

außerdem

$$|(\bar{\gamma}_1, \Gamma_1)' W (\bar{\gamma}_1, \Gamma_1)| = |(\bar{\gamma}_1, \Gamma_1)|^2 |W|$$

$$\log |(\bar{\gamma}_1, \Gamma_1)' W (\bar{\gamma}_1, \Gamma_1)| = \frac{1}{2} \log |(\bar{\gamma}_1, \Gamma_1)' W (\bar{\gamma}_1, \Gamma_1)| - \frac{1}{2} \log |W|$$

$$(2.19) \quad = \frac{1}{2} \log |\bar{\gamma}_1' W \bar{\gamma}_1| - \frac{1}{2} \log |W|$$

(2.18) und (2.19) sind also die Werte, für die die Likelihood-Funktion (2.15) in Bezug auf A_1^* maximiert wird. Einsetzen von (2.18) und (2.19) ergibt nun die konzentrierte Likelihood-Funktion

$$(2.20) \quad L^*(\dots) = c_3 + \frac{1}{2} \log |\bar{\gamma}_1' W \bar{\gamma}_1| - \frac{1}{2} \log |W| - \frac{1}{2} \log \sigma_{11} \\ - \frac{1}{2} \sigma_{11}^{-1} \bar{\alpha}_1' M \bar{\alpha}_1$$

Die Herleitung der konzentrierten Likelihood-Funktion (2.20) ist als Spezialfall in der Herleitung von Dhrymes¹⁾ enthalten, wo sie nicht für eine Gleichung, sondern für ein Untersystem hergeleitet wird.

L^* aus (2.20) wird nun weiter konzentriert, indem L bezüglich σ_{11} maximiert wird:

$$\frac{\partial L^*}{\partial \sigma_{11}} = - \frac{1}{2} \sigma_{11}^{-1} + \sigma_{11}^{-2} \bar{\alpha}_1' M \bar{\alpha}_1$$

Nullsetzen $\frac{\partial L^*}{\partial \sigma_{11}} = 0$ resultiert in

$$(2.21) \quad \sigma_{11}(\bar{\alpha}_1) = \bar{\alpha}_1' M \bar{\alpha}_1.$$

1) vgl. Dhrymes, P.J.: Econometrics, a.a.O. S.330ff..

Einsetzen von $\sigma_{11}(\bar{\alpha}_1)$ in (2.20) ergibt die zweifach konzentrierte Likelihood-Funktion

$$(2.22) \quad L^{**}(\bar{\alpha}_1) = c_4 + \frac{1}{2} \log(\bar{\gamma}_1' W \bar{\gamma}_1) - \frac{1}{2} \log(\bar{\alpha}_1' M \bar{\alpha}_1)$$

Jetzt werden Nullrestriktionen in $\bar{\gamma}_1$ und $\bar{\alpha}_1$ berücksichtigt (nicht jedoch eine Normierung) und wieder allgemein L^{**} für die i-te Strukturgleichung betrachtet:

$$(2.22') \quad L^{**}(\bar{\alpha}_i) = c_4 + \frac{1}{2} \log(\bar{\gamma}_i' W_{ii} \bar{\gamma}_i) - \frac{1}{2} \log(\bar{\alpha}_i' M_{ii} \bar{\alpha}_i)$$

$$\text{mit } W_{ii} = \frac{1}{T} \bar{Y}_i' (I - X(X'X)^{-1}X') \bar{Y}_i$$

$$M_{ii} = \frac{1}{T} \bar{Z}_i' \bar{Z}_i$$

In $\bar{Y}_i$ und $\bar{Z}_i$ treten also im Gegensatz zu Y_i und Z_i auch die Beobachtungen der Abhängigen der i-ten Gleichung auf.

Maximieren von $L^{**}(\bar{\alpha}_i)$ bzgl. β_i bedeutet differenzieren

$$\text{von } \bar{\alpha}_i' M_{ii} \bar{\alpha}_i = \frac{1}{T} (\bar{\gamma}_i', \beta_i') \begin{pmatrix} \bar{Y}_i' \bar{Y}_i & \bar{Y}_i' X_i \\ X_i' \bar{Y}_i & X_i' X_i \end{pmatrix} \begin{pmatrix} \bar{\delta}_i \\ \beta_i \end{pmatrix} \text{ nach } \beta_i \text{ und Nullsetzen.}$$

$$\frac{\partial \bar{\alpha}_i' M_{ii} \bar{\alpha}_i}{\partial \beta_i} = \frac{1}{T} \frac{\partial}{\partial \beta_i} (2\beta_i' X_i' \bar{Y}_i \bar{\delta}_i + \beta_i' X_i' X_i \beta_i) = 0 \quad \Rightarrow$$

$$(2.23) \quad \beta_i(\bar{\delta}_i) = -(X_i' X_i)^{-1} X_i' \bar{Y}_i \bar{\delta}_i$$

Durch einsetzen von $\beta_i(\bar{\delta}_i)$ in (2.22') erhält man schließlich die dreifach konzentrierte (Log-) Likelihood-Funktion

$$\begin{aligned} L^{***}(\bar{\delta}_i) &= c_5 + \frac{1}{2} \log(\bar{\delta}_i' W_{ii} \bar{\delta}_i) - \frac{1}{2} \log(\bar{\delta}_i' W_{ii}^* \bar{\delta}_i) \\ (2.24) \quad &= c_5 - \frac{1}{2} \log \left(\frac{\bar{\delta}_i' W_{ii}^* \bar{\delta}_i}{\bar{\delta}_i' W_{ii} \bar{\delta}_i} \right) \end{aligned}$$

$$\text{mit } W_{ii}^* = \bar{Y}_i' (I - X_i (X_i' X_i)^{-1} X_i') \bar{Y}_i \quad .$$

Um zu (LIML-) Schätzwerten für $\bar{\gamma}_1$ zu kommen, ist (2.24) zu maximieren. $L^{**}(\bar{\gamma}_1)$ ist jedoch homogen vom Grade 0 in $\bar{\gamma}_1$, so daß man, um zu eindeutigen Schätzwerten zu gelangen, $\bar{\gamma}_1$ normieren muß. Dies geschieht der Einfachheit halber durch die Bedingung

$$(2.25) \quad \bar{\gamma}_1' W_{11} \bar{\gamma}_1 = c \quad \text{mit einem beliebigen } c > 0.$$

Maximieren von (2.24) bedeutet nun offensichtlich minimieren des Ausdrucks

$$(2.26) \quad \lambda(\bar{\gamma}_1) = \frac{\bar{\gamma}_1' W_{11}^* \bar{\gamma}_1}{\bar{\gamma}_1' W_{11} \bar{\gamma}_1} \quad \text{unter der Nebenbedingung (2.25).}$$

Zum Minimieren von Funktionen unter Nebenbedingungen ist das Lagrangesche Multiplikatorenverfahren das geeignete Verfahren: Zu minimieren ist die Lagrange-Funktion

$$f(\bar{\gamma}_1, \lambda) = \bar{\gamma}_1' W_{11}^* \bar{\gamma}_1 - \lambda(\bar{\gamma}_1' W_{11} \bar{\gamma}_1 - c)$$

bezüglich $\bar{\gamma}_1$ und des Lagrange-Multiplikators λ .

Es ist

$$\frac{\partial f}{\partial \lambda} = \bar{\gamma}_1' W_{11} \bar{\gamma}_1 - c = 0$$

$$\frac{\partial f}{\partial \bar{\gamma}_1} = 2(W_{11}^* - W_{11}) \bar{\gamma}_1 = 0$$

d.h. der Schätzwert von $\bar{\gamma}_1$ wird bestimmt als Lösung des homogenen Gleichungssystems

$$(2.27) \quad (W_{11}^* - \lambda W_{11}) \bar{\gamma}_1 = 0$$

Dieses besitzt nur dann eine nichttriviale Lösung, wenn

$$|(W_{11}^* - \lambda W_{11})| = 0,$$

bzw. wenn

$$(2.28) \quad |(D' W_{11}^* D - I)| = 0$$

(D ist die Matrix, für die $D'W_{ii}D = I$. Eine solche Matrix existiert wegen der positiven Definitheit von W_{ii} immer.) (2.28) ist ein Eigenwertproblem mit einer positiv definiten Matrix $D'W_{ii}^*D$, das daher nur reelle und positive Wurzeln besitzt. Welche Wurzel ist nun zu wählen, um (2.26) zu minimieren? Offensichtlich ist das λ aus (2.27) gleich dem λ aus (2.26) (wie man durch Prämultiplikation mit $\bar{\delta}_i'$ sieht). Das heißt aber, daß die kleinste Wurzel von (2.28) zu wählen ist. Auflösen des Gleichungssystems (2.27) mit diesem λ_1 ergibt dann einen eindeutigen Schätzwert $\hat{\alpha}_i$ für $\bar{\delta}_i$. Einsetzen dieses $\hat{\alpha}_i$ in (2.23) ergibt den LIML-Schätzwert $\hat{\beta}_1$ für β_1 .

Zusammenfassend läßt sich folgende Definition formulieren:

Definition 2.3: Es sei

$$\bar{Z}_1 \bar{\alpha}_1 + u_1 = 0 = \bar{Y}_1 \bar{\delta}_1 + X_1 \beta_1 + u_1$$

die i-te Strukturgleichung in der homogenen Form (d.h. in $\bar{Z}_1$ befindet sich auch die abhängige Variable der i-ten Gleichung). Die Schätzfunktion $\hat{\alpha}_1$, die man als Lösung der Gleichungssysteme

$$(W_{ii}^* - \lambda W_{ii}) \bar{\delta}_i = 0$$

und $\beta_1(\bar{\delta}_1) = -(X_1' X_1)^{-1} X_1' Y_1 \bar{\delta}_1$ unter der Normierung (2.25)

erhält, heißt LIML-Schätzfunktion von $\bar{\alpha}_1$. Dabei ist λ die kleinste Wurzel des zugehörigen Eigenwertproblems. W_{ii} bzw. W_{ii}^* sind in (2.22') bzw. in (2.24) definiert. Zur LIML-Schätzfunktion $\hat{\alpha}_1$ von $\bar{\alpha}_1$ aus (2.1) gelangt man, indem man $\hat{\alpha}_1$ durch $-\hat{\alpha}_{i1}$ dividiert. Dies ist wegen der freien Wahl des c in der Normierungsbedingung (2.25) möglich.

Man kann $\hat{\alpha}_1$ aus Definition 2.3 auch ohne Annahme der Normalverteilung der Störglieder herleiten. Man spricht dann von einem Quasi-Maximum-Likelihood-Verfahren.

2.3.2. Die LVR-Methode

Wie der vorangehende Abschnitt gezeigt hat, ist die Herleitung der LIML-Schätzfunktion sehr mühsam. Nun läßt sich jedoch das LISE-Schätzverfahren auch direkt aus dem (letztlich zu minimierenden) Ausdruck (2.26) begründen, wenn man ihn als Verhältnis zweier (unten noch genauer zu erklärenden) Varianzen auffaßt. Die danach durch Minimierung als LVR (least variance ratio/kleinstes Varianzverhältnis) bestimmte Schätzfunktion stimmt natürlich mit der durch das Maximum-Likelihood-Prinzip begründeten LIML-Schätzfunktion überein.

Die zu minimierende Funktion (2.26) lautet

$$(2.29) \quad \lambda(\bar{y}_1) = \frac{\bar{y}_1' W_{11}^* \bar{y}_1}{\bar{y}_1' W_{11} \bar{y}_1} \quad \text{mit}$$

$$W_{11} = T^{-1} \bar{Y}_1' (I - X(X'X)^{-1}X') \bar{Y}_1$$

$$W_{11}^* = T^{-1} \bar{Y}_1' (I - X_1(X_1'X_1)^{-1}X_1') \bar{Y}_1.$$

Wie eine einfache Rechnung zeigt, ist der Zähler von (2.29) nichts anderes als die OLS-Reststreuung $T^{-1} \hat{u}_1' \hat{u}_1$ des Ansatzes¹⁾

$$(2.30) \quad -\bar{Y}_1 \bar{y}_1 = X_1 \beta_1 + u_1^*,$$

während der Nenner von (2.29) die OLS-Reststreuung $T^{-1} \hat{u}_1' \hat{u}_1$ des Ansatzes

$$(2.31) \quad -\bar{Y}_1 \bar{y}_1 = X \beta + u_1$$

ist²⁾. In (2.30) sind alle Nullrestriktionen bzgl. gemein-

1) vgl. Schönfeld, P.: Methoden der Ökonometrie, Bd.II, a.a.O. S.262.

2) Es ist $\hat{u}_1 = \bar{Y}_1 \bar{y}_1 - \bar{Y}_1 \hat{\beta}_1 = \bar{Y}_1 \bar{y}_1 - X_1 \hat{\beta}_1 = \bar{Y}_1 \bar{y}_1 - X_1 (X_1' X_1)^{-1} X_1' \bar{Y}_1 \bar{y}_1 = (I - X_1 (X_1' X_1)^{-1} X_1') \bar{Y}_1 \bar{y}_1$. Ebenso berechnet man $\hat{u}_1$.

sam abhängiger und vorherbestimmter Variablen berücksichtigt, in (2.31) sind keine Vorherbestimmten ausgeschlossen. Wegen der im Vergleich zu (2.31) zusätzlichen Restriktionen in (2.30) ist der Zähler von (2.29) immer mindestens so groß wie der Nenner. Sie sind genau dann gleich ($\lambda=1$), wenn die a-priori Nullrestriktionen in β_1 korrekt sind. Minimieren von (2.29) bedeutet also, die Parameter der i-ten Strukturgleichung so zu bestimmen, daß die a-priori Restriktionen bezüglich der Vorherbestimmten möglichst gut mit dem Beobachtungsmaterial (repräsentiert durch X_i und $\bar{Y}_i$) in Einklang sind.

2.3.3. Die LI/LGRV-Methode

Eine weitere Möglichkeit, die LISE-Schätzungen zu motivieren und herzuleiten, ist die LI/LGRV (limited information/least generalized residual variance/ kleinste verallgemeinerte Reststreuung bei begrenzter Information)-Methode. Unter der verallgemeinerten Reststreuung der reduzierten Form versteht man den Ausdruck

$$|T^{-1}V \cdot V|$$

Man kann zeigen, daß minimieren dieser Determinante gerade die OLS-Schätzung für Π ergibt¹⁾.

Betrachtet man nur den bezüglich der i-ten Strukturgleichung relevanten Teil der reduzierten Form

$$Y_i = X_i \Pi_i + V_i,$$

so würde man durch minimieren von

$$(2.32) \quad |T^{-1}V_i V_i|$$

eine OLS-Schätzung für Π_i bekommen. Von dieser Schätzung

1) vgl. Schönfeld, P.: Methoden der Ökonometrie, Bd.II, a.a.O. S.265.

könnte man jedoch nur dann eindeutig auf die Strukturparameter schließen, wenn das Gleichungssystem $\pi_{2i}\bar{y}_i = 0$ eine nichttriviale Lösung besitzt (vgl. I, (3.3')). Dies erzwingt man durch die Nebenbedingungen

$$(2.33) \quad \pi_{2i}\bar{y}_i = 0 \quad \text{und} \quad \bar{y}_i' w_{ii} \bar{y}_i = c \quad (c > 0).$$

Beim LI/LGRV-Verfahren minimiert man also den Logarithmus des Ausdrucks (2.32) unter den Nebenbedingungen (2.33) mit Hilfe der Lagrangeschen Multiplikatoren durch differenzieren nach π_i , $\bar{y}_i$ und den Multiplikatoren. Man erhält dann zunächst eine Schätzung für π_i in Abhängigkeit von $\bar{y}_i$ und schließlich die Schätzgleichungen für $\bar{y}_i$, die genau denen von (2.27), also den LIML-Schätzgleichungen, entsprechen: LIML-, LVR-, LI/LGRV-Schätzfunktionen stimmen überein. Auf eine genaue Herleitung der LISE-Schätzungen nach dem LI/LGRV-Prinzip wird hier verzichtet. Sie ist z.B. bei Goldberger oder Schönfeld zu finden¹⁾.

2.3.4. LISE als IV-Schätzung

Ebenso wie 2SLS lassen sich auch die LISE-Schätzungen als IV-Schätzung interpretieren, und zwar unter der Normierung $\bar{y}_{ii} = -1$ als IV-Schätzung mit der Instrumentenmatrix

$$(2.34) \quad P_i = ((I - \lambda \bar{M})Y_i, X_i) .$$

Dabei ist $\bar{M} = (I - X(X'X)^{-1}X')$ und λ das λ aus Definition 2.3, also die kleinste Wurzel des Eigenwertproblems (2.27).

Zunächst wird die formale Übereinstimmung der IV-Schätzfunktion für α_i (die Parameter der i-ten Strukturgleichung)

1) vgl. Goldberger, A.S.: *Econometric Theory*, a.a.O. S.338ff.

Schönfeld, P.: *Methoden der Ökonometrie*, Bd.II, a.a.O. S.266ff.

mit obiger Instrumentenmatrix P_1 aus (2.34) mit der LISE-Schätzfunktion aus Definition 2.3 gezeigt. Die IV-Schätzfunktion lautet:

$$\hat{\alpha}_1 = (P_1' Z_1)^{-1} P_1' y_1 \quad \text{mit } P_1 \text{ aus (2.34)}$$

$$\Leftrightarrow P_1' Z_1 \hat{\alpha}_1 - P_1' y_1 = 0.$$

Läßt man nun die Normierung $\bar{\delta}_{11} = -1$ wieder fallen, so erhält man

$$(2.35) \quad \bar{P}_1' \bar{Z}_1 \hat{\alpha}_1 = 0 \quad \text{mit } \bar{P}_1 = ((I - \lambda M) \bar{Y}_1, X_1).$$

$\bar{Y}_1$ setzt sich aus Y_1 und y_1 zusammen.

Aufspalten von (2.35) ergibt

$$\bar{P}_1' \bar{Y}_1 \bar{\delta}_1 + \bar{P}_1' X_1 \beta_1 = 0 \quad \text{oder}$$

$$(2.36) \quad \begin{aligned} \bar{Y}_1' (I - \lambda M) \bar{Y}_1 \bar{\delta}_1 + \bar{Y}_1' X_1 \beta_1 &= 0 \\ X_1' \bar{Y}_1 \bar{\delta}_1 + X_1' X_1 \beta_1 &= 0 \end{aligned}$$

Auflösen nach β_1 resultiert in

$$(2.37) \quad \beta_1(\bar{\delta}_1) = -(X_1' X_1)^{-1} X_1' \bar{Y}_1 \bar{\delta}_1$$

Dies ist genau das β_1 aus Definition 2.3. Einsetzen von (2.37) in (2.36) ergibt

$$(\bar{Y}_1' (I - X_1 (X_1' X_1)^{-1} X_1') \bar{Y}_1 - \lambda \bar{Y}_1' M \bar{Y}_1) \bar{\delta}_1 = 0 \quad \text{oder}$$

$$(w_{11}^* - \lambda w_{11}) \bar{\delta}_1 = 0 \quad \text{mit den } w_{11}^* \text{ und } w_{11} \text{ aus Defi-}$$

nition 2.3. Die LISE-Schätzfunktion aus Definition 2.3 läßt sich also in der Tat zumindest formal als IV-Schätzfunktion interpretieren.

Sind nun aber auch die Bedingungen (1.5) und (1.6), nämlich

$$\text{plim } T^{-1}P_i u_i = 0 \quad \text{und}$$

$$\text{plim } T^{-1}P_i z_i = M_i \quad (|M_i| \neq 0)$$

mit dem P_i aus (2.34) erfüllt? Aus (2.34) erkennt man leicht, daß sich diese Frage auf die entsprechende Frage für die 2SLS-Schätzfunktion (vgl. 2.2.1.) reduzieren läßt, falls man nachweisen kann, daß

$$\text{plim } \hat{\lambda} = 1 \quad .$$

Dies ist in der Tat der Fall¹⁾. Mit W_{ii}^* und W_{ii} aus Definition 2.3 und $\bar{M} = I - X(X'X)^{-1}X'$, $\bar{M}_i = I - X_i(X_i'X_i)^{-1}X_i'$ ist

$$\begin{aligned} \text{plim } (W_{ii}^* - W_{ii}) &= \text{plim } T^{-1} \bar{Y}_i' (\bar{M}_i - \bar{M}) \bar{Y}_i \\ &= \text{plim } T^{-1} \Pi_{2i}' X_i^* \bar{M}_i X_i^* \Pi_{2i} \quad , \text{ wo} \end{aligned}$$

Π_{2i} der Untermatrix Π_{2i} von Π aus I, Satz 3.2 entspricht, und X_i^* die aus der i -ten Gleichung ausgeschlossenen vorherbestimmten Variablen beinhaltet. Wegen I, Annahme 2(a) existiert auch $\text{plim } T^{-1} X_i^* \bar{M}_i X_i^* = M_{ii}$, so daß

$$\begin{aligned} \text{plim } |W_{ii}^* - \hat{\lambda} W_{ii}| &= |\text{plim}(W_{ii}^* - W_{ii}) - \text{plim}(\hat{\lambda} - 1) \text{plim } W_{ii}| \\ (2.38) \quad &= |\Pi_{2i}' M_{ii} \Pi_{2i} - \text{plim}(\hat{\lambda} - 1) \Omega_{ii}|. \end{aligned}$$

$\text{plim } W_{ii} = \Omega_{ii}$, eine Untermatrix von Ω , der reduzierten Form-Kovarianzmatrix, ist positiv definit. Nun ist wegen der Rangbedingung für die Identifizierbarkeit (vgl. I, Satz 3.2 : $\text{rg}(\Pi_{2i}) = N_i - 1$) die Matrix $\Pi_{2i}' M_{ii} \Pi_{2i}$ singulär.

1) Die Berechnung der plim's erfolgt wie gewohnt: $\bar{Y}_i$ wird durch den entsprechenden reduzierten-Form-Anteil ersetzt und der plim direkt angewandt unter Beachtung der stochastischen Annahmen in Kapitel I.

Das heißt jedoch, daß die kleinste Wurzel von (2.38) Null ist und damit $\text{plim}(\lambda - 1) = 0$, d.h. $\text{plim}\lambda = 1$.

Aus dem oben Gesagten in Verbindung mit Satz 1.4 erhält man

Satz 2.3: Die LISE (LIML-, LVR-, LI/LGRV-) -Schätzfunktion $\hat{\alpha}_1$ aus Definition 2.3 ist eine konsistente Schätzfunktion für α_1 aus Gleichung (2.1).

2.4. Die k-Klasse

Wie bereits zu Beginn von Abschnitt 2 erwähnt, lassen sich alle Einzelgleichungsschätzverfahren in einer Klasse, der sogenannten k-Klasse¹⁾ zusammenfassen. Die genaue Definition der k-Klasse und der Beweis der obigen Aussage folgt in den nächsten beiden Unterabschnitten. Zwei einfache Verallgemeinerungen der k-Klasse werden im dritten Unterabschnitt behandelt.

2.4.1. Herleitung und Definition der k-Klasse (IV-Interpretation der k-Klasse)

Aus Definition 2.2 und der Diskussion in 2.3.4. liegt die Definition der k-Klasse recht nahe. Denn sowohl 2SLS als auch LIML sollen zur k-Klasse gehören. Läßt man in (2.34), der Definition der Instrumentenmatrix für die LISE (LIML-) -Schätzung, statt λ ein beliebiges skalares k , das stochastisch oder nichtstochastisch sein darf, zu, so erhält man folgende Definition der k-Klasse:

1) vgl. Theil, H.: Economic Forecasts and Policy, 2nd ed., Amsterdam 1961, S.231.

Definition 2.4: Die einparametrische Schar von Schätzfunktionen für α_1 aus Gleichung (2.1), die durch die Schätzgleichungen

$$(2.39) \quad Z_1'(I - k\bar{M})Z_1\alpha_1(k) = Z_1'(I - k\bar{M})y_1$$

bestimmt ist, heißt k -Klasse. Dabei ist der Parameter k ein beliebiger stochastischer oder nichtstochastischer Skalar, $\bar{M} = I - X(X'X)^{-1}X'$.

Eine etwas suggestivere Motivation der k -Klasse ist die folgende: Wie man leicht nachrechnet, lassen sich die 2SLS-Schätzgleichungen (vgl. Def. 2.2) in der folgenden Weise darstellen:

$$(2.40) \quad \begin{pmatrix} Y_1'Y_1 - \hat{V}_1'\hat{V}_1 & Y_1'X_1 \\ X_1'Y_1 & X_1'X_1 \end{pmatrix} \begin{pmatrix} \hat{\delta}_1 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} Y_1' - \hat{V}_1' \\ X_1' \end{pmatrix} y_1,$$

wo $\hat{V}_1 = Y_1 - \hat{Y}_1 = (I - X(X'X)^{-1}X')Y_1$ der OLS-geschätzte Anteil der reduzierten-Form-Residuen ist, der zu Y_1 gehört.

(2.40) zeigt, daß sich 2SLS nur dadurch von OLS unterscheidet, daß von $Y_1'Y_1$ noch $\hat{V}_1'\hat{V}_1$ subtrahiert wird. Dhrymes¹⁾

spricht davon, daß die Y_1 , die erklärenden abhängigen Variablen der i -ten Gleichung, gewissermaßen von dem (geschätzten) stochastischen Anteil bereinigt würden, der mit dem Störglied der i -ten Gleichung korreliert ist.

Diese Bereinigung bewirkt dann die günstigen stochastischen Eigenschaften von 2SLS (insbesondere also die Konsistenz).

Von den 2SLS-Schätzgleichungen in der Form (2.40) kommt man zu den k -Klasse-Schätzgleichungen, indem man $Y_1'Y_1$ nicht um $\hat{V}_1'\hat{V}_1$ sondern um einen beliebigen proportionalen

1) vgl. Dhrymes, P.J.: Econometrics, a.a.O. S.200.

Anteil davon bereinigt. Die (2.40) entsprechende Darstellung der k-Klasse-Schätzgleichungen lautet also:

$$(2.41) \quad \begin{pmatrix} Y_1' Y_1 - k \hat{V}_1' \hat{V}_1 & Y_1' X_1 \\ X_1' Y_1 & X_1' X_1 \end{pmatrix} \begin{pmatrix} \hat{\beta}_1(k) \\ \hat{\beta}_1(k) \end{pmatrix} = \begin{pmatrix} Y_1' - k \hat{V}_1' \\ X_1' \end{pmatrix} y_1$$

Wie zu Beginn dieses Abschnittes bereits angedeutet besitzt die k-Klasse (formal) eine IV-Interpretation mit der Instrumentenmatrix

$$(2.42) \quad P_1(k) = ((I - k\bar{M})Y_1, X_1), \text{ wo } \bar{M} = I - X(X'X)^{-1}X'.$$

Dies ergibt sich unmittelbar aus 2.3.4., wenn man in (2.34) $\hat{\lambda}$ durch k ersetzt. Als Klasse von IV-Schätzfunktionen kann die k-Klasse jedoch nur dann aufgefaßt werden, wenn sie die Bedingungen (1.5) und (1.6), die die Konsistenz implizieren, erfüllt. Diese Bedingungen werden aber nur von einer Unterklasse der k-Klasse erfüllt, wie folgender Satz zeigt:

Satz 2.4: $\mathcal{M}$ sei die Menge der in Definition 2.4 zugelassenen k.

$\mathcal{M}_1$ sei die Untermenge von $\mathcal{M}$, für deren Elemente gilt: $\text{plim}(k-1) = 0$.

$\mathcal{M}_2$ sei die Untermenge von $\mathcal{M}$, deren Elemente durch die Schätzgleichungen (2.39) eine konsistente Schätzfunktion bestimmen.

Dann gilt

$$\mathcal{M}_1 = \mathcal{M}_2$$

Beweis: $\mathcal{M}_1 = \mathcal{M}_2$ genau dann, wenn man zeigen kann, daß jedes Element aus $\mathcal{M}_1$ auch Element von $\mathcal{M}_2$ und umgekehrt jedes Element aus $\mathcal{M}_2$ auch Element von $\mathcal{M}_1$ ist.

1. Sei also $k \in \mathcal{M}_1$. Dann gilt (mit dem P_1 aus (2.42))

$$(2.43) \quad \text{plim} T^{-1} P_i'(k') Z_i = \begin{pmatrix} \Pi_i' M \Pi_i + \text{plim}(1-k') D_i' \Sigma D_i & , & \Pi_i' M_i \\ M_i' \Pi_i & , & M_{ii} \end{pmatrix}$$

Dies ergibt sich durch ersetzen von Y_i durch $XW_i + V_i$ und direkte Anwendung des plim's unter Beachtung von I, Annahme 2(a) und Folgerung 1.

$$(2.44) \quad \text{plim} T^{-1} P_i'(k') u_i = \begin{pmatrix} \text{plim}(1-k') D_i' \sigma_i \\ 0 \end{pmatrix}.$$

Auf eine genaue Herleitung von (2.43) und (2.44) kann hier verzichtet werden, da sie in Analogie zum Beweis von Satz 1.2 erfolgt, wo das Verfahren ausführlich dargestellt wurde. Da $k \in \mathcal{M}_1$ ist $\text{plim}(1-k')=0$ und der plim aus (2.43) existiert als nichtsinguläre Matrix, während (2.44) Null wird. Das bedeutet, daß die beiden Bedingungen (1.5) und (1.6) für eine IV-Schätzung erfüllt sind und die k -Klasse mit $k \in \mathcal{M}_1$ konsistent ist, d.h. $k \in \mathcal{M}_2$.

2. Sei umgekehrt $k^* \in \mathcal{M}_2$, d.h. $\hat{\alpha}_1(k^*)$ ist konsistent, und es gilt

$$(2.45) \quad \text{plim} \hat{\alpha}_1(k^*) = \alpha_1 + (\text{plim} T^{-1} P_i'(k^*) Z_i)^{-1} (\text{plim} T^{-1} P_i'(k^*) u_i) \\ = \alpha_1 \quad \text{wegen der Konsistenz.}$$

Der zweite Summand auf der rechten Seite von (2.45) muß also Null sein. Dies ist nur dann der Fall, wenn

$\text{plim} T^{-1} P_i'(k^*) Z_i$ als nichtsinguläre Matrix existiert und $\text{plim} T^{-1} P_i'(k^*) u_i$ gleich Null ist. Wie (2.43) und (2.44) zeigen, ist dies genau dann der Fall, wenn $\text{plim}(k^*-1)$ gleich Null ist. Dies heißt jedoch, daß $k^* \in \mathcal{M}_1$.

q.e.d.

2.4.2. OLS, ILS, 2SLS, LISE als k-Klasse-Schätzungen

Nach dem Vorangegangenen sieht man nun leicht, daß alle bisher behandelten Einzelgleichungsschätzverfahren (außer GLS) zur k-Klasse gehören. Es läßt sich zusammenfassend folgender Satz formulieren:

Satz 2.5: Aus den k-Klasse-Schätzgleichungen (2.39) erhält man bei verschiedener Wahl von k die folgenden Schätzfunktionen für α_1 aus Gleichung (2.1): Für

- a) $k = 0$ die OLS-Schätzfunktion
- b) $k = 1$ die 2SLS-Schätzfunktion
- c) $k = \lambda$ die LISE (LIML, LVR, LI/LGRV)-Schätzfunktion¹⁾
(λ ist die kleinste Wurzel des Eigenwertproblems (2.28))
- d) Im genau identifizierten Fall der i-ten Strukturgleichung erhält man außerdem für
 $k = 1$ die ILS-Schätzfunktion.

Beweis: a) und b) beweist man leicht durch einsetzen von 0 bzw. 1 in die Schätzgleichungen (2.39) und Vergleich mit (1.2) bzw. (2.3).

c) erhält man durch Vergleich der beiden Instrumentenmatrizen (2.34) und (2.42).

d) zeigt man, indem man nachweist, daß im genau identifizierten Fall die 2SLS- mit der ILS-Schätzfunktion übereinstimmt:

Nach Definition 2.1 ist die ILS-Schätzfunktion durch folgende Schätzgleichungen bestimmt:

$$\begin{pmatrix} \hat{w}_{11} & I \\ \hat{w}_{21} & 0 \end{pmatrix} \begin{pmatrix} \hat{\delta}_1 \\ \beta_1 \end{pmatrix} = 0 \quad \text{bzw.}$$

1) Genau genommen gilt dies nur asymptotisch (vgl. Schönfeld, P.: Methoden der Ökonometrie, Bd.II, a.a.O. S.276).

$$\hat{\Pi}_{1\hat{y}_1} = \begin{pmatrix} \hat{\beta}_1 \\ 0 \end{pmatrix}$$

oder wegen $\hat{\Pi}_1 = (X'X)^{-1}X'\bar{Y}_1$ und $\bar{Y}_1 = (y_1, y_1)$ unter der Normierung $\bar{y}_{11} = -1$

$$-(X'X)^{-1}X'y_1 + (X'X)^{-1}X'y_1\hat{\beta}_1 = -\hat{\beta}_1.$$

Prämultiplikation mit $Z_1'X$ ergibt

$$-Z_1'X(X'X)^{-1}X'y_1 + Z_1'X(X'X)^{-1}X'y_1\hat{\beta}_1 = -Z_1'X\hat{\beta}_1.$$

Zusammenfassend:

$$Z_1'X(X'X)^{-1}X'Z_1\hat{\alpha}_1 = Z_1'X(X'X)^{-1}X'y_1.$$

Das sind genau die 2SLS-Schätzgleichungen.

Umgekehrt ist wegen der genauen Identifizierbarkeit $\bar{R}_1 = N_1 - 1$ oder $R = N_1 + R_1 - 1$. Das heißt aber, daß $Z_1'X$ quadratisch und i.a. nichtsingulär ist. Man kann also auch von den 2SLS-Schätzgleichungen ausgehend durch Prämultiplikation mit $(Z_1'X)^{-1}$ zu den ILS-Schätzgleichungen gelangen.

q.e.d.

Auf Grund von Satz 2.5 und Satz 2.4 zeigt man nun leicht die Konsistenz bzw. Inkonsistenz der einzelnen Schätzfunktionen (die schon auf andere Weise hergeleitet wurde): OLS ist im allgemeinen Fall inkonsistent, da $\text{plim } k = \text{plim } 0 = 0 \neq 1$. ILS, 2SLS und LISE sind konsistent, da $\text{plim } 1 = \lim 1 = 1$ bzw. $\text{plim } \hat{\lambda} = 1$, wie in 2.3.4. nachgewiesen.

2.4.3. h- und Doppel-k-Klasse

Neben der k-Klasse haben zwei weitere Klassen von Einzelgleichungsschätzfunktionen, die eng mit der k-Klasse zusammenhängen, eine gewisse Bedeutung erlangt, nämlich die h- und die Doppel-k-Klasse, die von Theil bzw. Nagar¹⁾ eingeführt wurden. Zur Herleitung der h-Klasse geht man wieder von der i-ten Strukturgleichung (2.1) aus. Man bereinigt nun Y_i durch einen proportionalen Anteil der für Y_i relevanten reduzierten-Form-Residuen $\hat{v}_i$ (die OLS-geschätzt sind), so daß man aus (2.1) erhält:

$$(2.46) \quad y_i = (Y_i - h\hat{v}_i) \gamma_i + X_i \beta_i + u_i^*$$

OLS-Schätzung von (2.46) ergibt die Schätzgleichungen für die Schätzfunktionen der h-Klasse. Auch hier ist h wieder ein beliebiger stochastischer oder nichtstochastischer Skalar.

$$(2.47) \quad \begin{pmatrix} Y_i' Y_i - h(2-h) \hat{v}_i' \hat{v}_i & Y_i' X_i \\ X_i' Y_i & X_i' X_i \end{pmatrix} \alpha_i(h) = \begin{pmatrix} Y_i' - h \hat{v}_i' \\ X_i' \end{pmatrix} y_i$$

Macht man in (2.47) die Substitution $h \rightarrow 1-h^*$, so erhält man die h-Klasse-Schätzgleichungen in "symmetrischer" Form, wo auf der linken Seite von (2.47) $(1-h^*)$, auf der rechten Seite $(1-h^*)$ als Proportionalitätsfaktor vor $\hat{v}_i' \hat{v}_i$ bzw. $\hat{v}_i'$ steht.

Bei der Doppel-k-Klasse geht man noch einen Schritt weiter: Rechts und links in (2.47) läßt man verschiedene voneinander unabhängige Proportionalitätsfaktoren k_1 und k_2 zu. Für $k_1=k_2$ reduziert sich die Doppel-k-Klasse dann natürlich wieder auf die einfache k-Klasse. Aus dem oben Gesagten bekommt man, wenn man wieder beachtet, daß

$$\hat{v}_i = (I - X(X'X)^{-1}X')Y_i = \bar{M}Y_i, \text{ die folgende Definition für}$$

1) Theil, H.: Economic Forecasts and Policy, 2nd ed., a.a.O. S.353f.

Nagar, A.L.: Double k-Class Estimators of Parameters in Simultaneous Equations and their Small Sample Properties, International Economic Review 3, 1962, 168-188.

die h- bzw. Doppel-k-Klasse:

Definition 2.5: a) Die einparametrische Schar von Schätzfunktionen für α_1 aus (2.1), die durch die Schätzgleichungen

$$(2.48) \quad Z_1'(I - (1-h^2)\bar{M})Z_1\hat{\alpha}_1(h) = Z_1'(I - (1-h)\bar{M})y_1$$

bestimmt ist, heißt h-Klasse.

b) Die zweiparametrische Schar von Schätzfunktionen für α_1 aus (2.1), die durch die Schätzgleichungen

$$(2.49) \quad Z_1'(I - k_1\bar{M})Z_1\hat{\alpha}_1(k_1, k_2) = Z_1'(I - k_2\bar{M})y_1$$

bestimmt ist, heißt Doppel-k-Klasse.

h , k_1 und k_2 sind dabei beliebige stochastische oder nicht-stochastische Skalare.

Die k-Klasse ist offensichtlich eine Unterklasse der Doppel-k-Klasse. Die Doppel-k-Klasse enthält damit alle Schätzfunktionen, die auch die k-Klasse enthält, nämlich OLS, 2SLS, LISE und ILS. Die h-Klasse ist ebenfalls eine Unterklasse der Doppel-k-Klasse (mit $k_1=1-h^2$, $k_2=1-h$), nicht jedoch der k-Klasse. Welche Schätzfunktionen haben k-Klasse und h-Klasse gemeinsam? Dies sind (siehe (2.48)) genau die Schätzfunktionen, für die $1-h^2=1-h$ oder

$$(2.50) \quad h(h-1) = 0$$

ist. Die beiden Lösungen von (2.50) sind $h=1$ und $h=0$.

Einsetzen in (2.48) ergibt die OLS- bzw. 2SLS- (im genau identifizierten Fall ILS-) Schätzfunktion.

Die Doppel-k-Klasse ist also die allgemeinste Klasse von Einzelgleichungsschätzverfahren. Wie man die Doppel-k-Klasse-Schätzfunktionen aus der 2SLS-Schätzfunktion konstruieren kann, zeigt Dhrymes¹⁾. Dieser Zusammenhang

1) Dhrymes, P.J.: Econometrics, a.a.O. S.359ff.

ergibt sich später (vgl. Abschnitt 5.1) als Spezialfall eines allgemeineren Zusammenhangs zwischen Doppel-K-Matrix-Klasse und 3SLS-Schätzfunktion.

Die Bedeutung von k - bzw. Doppel- k -Klasse liegt natürlich nicht nur darin, daß sie alle Einzelgleichungsschätzverfahren als Spezialfälle enthalten, sondern daß man wegen der Flexibilität in der Wahl von k bzw. k_1, k_2 beliebige andere Schätzfunktionen konstruieren kann. Insbesondere kann man diese Flexibilität in Bezug auf Klein-Stichprobeneigenschaften ausnutzen. Darauf wird noch in Kapitel III näher eingegangen.

2.5. Eigenschaften der Schätzungen der k -Klasse

Wie 2.4.3 gezeigt hat, ist die Doppel- k -Klasse die allgemeinste Klasse von Einzelgleichungsschätzfunktionen, die hier behandelten Einzelgleichungsschätzverfahren sind jedoch alle schon in der k -Klasse enthalten. Man kann sich also bei der Diskussion von Eigenschaften der Schätzfunktionen auf die k -Klasse beschränken. Dabei sind hier unter Schätzeigenschaften die - einfacher als Eigenschaften bei endlichem Stichprobenumfang zu behandelnden - asymptotischen Schätzeigenschaften zu verstehen. Auf Klein-Stichprobeneigenschaften wird in Kapitel III näher eingegangen.

Die wichtigste asymptotische Schätzeigenschaft, nämlich die Konsistenz, wurde bereits (auf relativ einfachem Wege) für ILS, 2SLS, LISE und allgemein für alle Schätzfunktionen der k -Klasse mit $\text{plim } k = 1$ nachgewiesen. Umgekehrt gibt es nach Satz 2.4 in der k -Klasse keine weiteren konsistenten Schätzfunktionen mehr¹⁾. Konsistenz einer Schätz-

1) Diese Aussagen sind immer als Aussagen im allgemeinen Fall, in dem die i -te Gleichung identifiziert ist, und erklärende abhängige Variablen tatsächlich auftreten, zu verstehen.

funktion für α_i aus (2.1) bedeutet, daß für unbegrenzt wachsenden Stichprobenumfang ($T \rightarrow \infty$) die Folge der Zufallsvariablen $\hat{\alpha}_i(T)$ gegen einen festen nichtstochastischen Wert, und zwar gegen den wahren Parameterwert α_i , konvergiert. Da jede Zufallsvariable eine Dichtefunktion besitzt, bedeutet dies, daß die Folge der zu den $\hat{\alpha}_i(T)$ gehörenden Dichtefunktionen sich in der Grenze in einem einzigen Punkt, nämlich α_i , konzentriert, die Folge der Dichtefunktionen also gegen eine Einpunktverteilung konvergiert. Wie sieht es nun jedoch mit der Güte der Konvergenz aus? Unter zwei konsistenten Schätzfunktionen würde man sicher diejenige wählen, bei der der Konzentrationsprozeß der Dichtefunktionen schon bei einem kleineren T als bei der anderen einsetzt. Für endliches T sind die Dichtefunktionen jedoch (bisher jedenfalls) nicht exakt bestimmbar. Um trotzdem die eventuell unterschiedliche Güte zweier konsistenter Schätzfunktionen beurteilen zu können, bedient man sich folgenden Verfahrens:

Man multipliziert die Folge der $\hat{\alpha}_i(T) - \alpha_i$ mit $\sqrt{T}$. Man kann dann zeigen, daß die Folge der zugehörigen Dichtefunktionen zwar nicht mehr gegen die Einpunktverteilung konvergiert, aber immer noch gegen die Dichtefunktion einer Verteilung mit Mittelwert Null und endlicher Kovarianzmatrix. Man hat also durch Multiplikation mit $\sqrt{T}$ die Dichtefunktionen der $\hat{\alpha}_i(T) - \alpha_i$ gewissermaßen "entzerrt". Unter zwei konsistenten Schätzfunktionen wird man dann diejenige wählen, bei der die Kovarianzmatrix der "asymptotischen Verteilung" (das ist die Verteilung, gegen die die Verteilungen der Folge der $\sqrt{T}(\hat{\alpha}_i(T) - \alpha_i)$ konvergieren) kleiner ist als bei der anderen¹⁾. Unter recht allgemeinen

1) Sind Σ_1 und Σ_2 zwei Kovarianzmatrizen, so ist Σ_1 kleiner als Σ_2 , wenn die Matrix $\Sigma_2 - \Sigma_1$ nicht-negativ-definit ist.

Voraussetzungen läßt sich sogar zeigen, daß die asymptotische Verteilung eine Normalverteilung ist, auch wenn man von den Störgliedern nicht annimmt, daß sie normalverteilt sind. Die Herleitung der asymptotischen Verteilungen in diesem allgemeinen Fall ist (besonders wenn verzögerte endogene Variablen unter den vorherbestimmten auftreten) sehr mühsam und würde den Rahmen dieser Arbeit sprengen. Eine exakte Herleitung, die auf der Anwendung von zentralen Grenzwertsätzen für sogenannte m -abhängige Variablen¹⁾ beruht, findet man bei Schönfeld²⁾. Die folgenden Sätze werden daher ohne Beweis wiedergegeben:

Satz 2.6: $\mathcal{M}$ sei die Menge der in Definition 2.4 zugelassenen k .

$\mathcal{M}'$ sei die Untermenge von $\mathcal{M}$, für deren Elemente gilt:
 $\text{plim } \sqrt{T}(k-1) = 0$.

$\mathcal{M}''$ sei die Untermenge von $\mathcal{M}$, deren Elemente durch die Schätzgleichungen (2.39) eine Schätzfunktion mit der asymptotischen Normalverteilung $N(\alpha_1, T^{-1}\sigma_{11}\bar{Q}_{11}^{-1})$ ³⁾ bestimmen. $\bar{Q}_{11}$ ist dabei die (nichtstochastische) Matrix

$$\text{plim } T^{-1}Z_1'X(X'X)^{-1}X'Z_1 = \begin{pmatrix} \Pi_1' M \Pi_1 & \Pi_1' M_1 \\ M_1' \Pi_1 & M_{11} \end{pmatrix} \quad (\text{vgl. (2.43)}).$$

Dann gilt:

$$\mathcal{M}' \subset \mathcal{M}'' \quad (\mathcal{M}' \text{ ist Untermenge von } \mathcal{M}'').$$

1) vgl. Schönfeld, P.: A Useful Central Limit Theorem for m -Dependent Variables, *Metrika* 17, 1971, Heft 1-2, S. 116-128.

2) vgl. Schönfeld, P.: Methoden der Ökonometrie, Bd. II, a.a.O.

3) Dies bedeutet also, daß $\sqrt{T}(\hat{\alpha}_1(T) - \alpha_1)$ asymptotisch $N(0, \sigma_{11}\bar{Q}_{11}^{-1})$ verteilt ist.

Man kann weiter zeigen, daß

$$\text{plim} \sqrt{T}(\lambda - 1) = 0$$

mit dem λ aus Definition 2.3. Außerdem ist auch $\text{plim} \sqrt{T}(1 - 1) = 0$. In Verbindung mit Satz 2.6 und Satz 2.5 erhält man so die

Folgerung: Die 2SLS- und LISE (LIML-, LVR-, LI/LGRV-) Schätzfunktionen (im genau identifizierten Fall auch die ILS-Schätzfunktion) für α_i aus (2.1) besitzen dieselbe asymptotische Normalverteilung, nämlich

$$N(\alpha_i, T^{-1} \sigma_{ii} \bar{Q}_{ii}^{-1}) ,$$

wo $\bar{Q}_{ii}$ in Satz 2.6 definiert ist.

2SLS und LISE besitzen dieselbe asymptotische Kovarianzmatrix $\sigma_{ii} \bar{Q}_{ii}^{-1}$, d.h. sie sind bezüglich ihrer asymptotischen Schätzeigenschaften nicht unterscheidbar und damit äquivalent. Man kann also, zumindest auf Grund der asymptotischen Schätzeigenschaften, nicht sagen, welche Schätzfunktion der anderen vorzuziehen ist.

Schätzverfahren, die auch konsistent sind, jedoch eine kleinere asymptotische Kovarianzmatrix als 2SLS und LISE besitzen, und damit diesen (zumindest theoretisch) vorzuziehen sind, sind die im nächsten Abschnitt zu behandelnden **f u l l - i n f o r m a t i o n**-Schätzverfahren (Schätzverfahren bei voller Information), die das Modell simultan schätzen und damit zur Schätzung einer Strukturgleichung die Information des gesamten Modells ausschöpfen.

3. Die Schätzverfahren der K-Matrix-Klasse

Ebenso wie sich die Einzelgleichungsschätzverfahren bei begrenzter Information in einer Klasse zusammenfassen lassen, so kann man auch die Schätzverfahren, die alle Parameter des Modells simultan schätzen und dabei die gesamte a-priori Information des Modells berücksichtigen (die sogenannten Simultanschätzverfahren bei voller Information (full information/FI)), in einer Klasse, der K-Matrix-Klasse, zusammenfassen¹⁾. Wie der Name schon sagt, wird die K-Matrix-Klasse durch Variation einer $N \times N$ -Parametermatrix erzeugt. Die Full-Information-Verfahren sind also Spezialfälle einer N^2 -parametrischen Schar von Schätzverfahren²⁾. Alle Limited-Information-Verfahren besitzen ein Full-Information-Pendant. Als Full-Information-Verallgemeinerung des Einzelgleichungs-OLS-Verfahrens läßt sich die GLS-Schätzung von Zellners Modell der scheinbar unverbundenen Regressionen (seemingly unrelated regressions)³⁾ - übertragen auf das interdependente Modell - betrachten, 2SLS besitzt in 3SLS und LIML in FIML seine FI-Verallgemeinerung. Wie sich noch zeigen wird, sind im genau identifizierten Fall 2SLS, 3SLS, LIML und FIML identisch mit der ILS-Methode. Die ILS-Methode läßt sich daher auch als FI-Methode betrachten, ohne daß man sie hier nochmals behandeln müßte. Schließlich erweist sich die K-Matrix-Klasse als FI-Verallgemeinerung der k-Klasse, und auch h- und Doppel-k-Klasse besitzen in h-Vektor- und Doppel-K-Matrix-Klasse ihr FI-Pendant.

1) vgl. Scharf, W.: K-Matrix-Class-Estimators..., a.a.O. 41-50.

2) Da schon die symmetrische K-Matrix-Klasse (also die Unterklasse der K-Matrix-Klasse mit symmetrischer Parametermatrix K) alle FI-Verfahren enthält, kann man sie sogar als Spezialfälle einer $N(N+1)/2$ -parametrischen Schar von Schätzverfahren betrachten.

3) vgl. Zellner, A.: An Efficient Method..., a.a.O. 348-368.

3.1. Die FI/OLS-Methode¹⁾

Statt einer einzigen Strukturgleichung betrachte man nun simultan alle N Strukturgleichungen des Gesamtmodells, wobei hier wie im Folgenden immer alle Gleichungen als identifizierbar vorausgesetzt seien:

$$(3.1) \quad y_i = Y_i \gamma_i + X_i \beta_i + u_i = Z_i \alpha_i + u_i \quad (i=1, \dots, N).$$

In Kurzschreibweise läßt sich (3.1) auch darstellen als

$$(3.2) \quad \mathbf{y} = \mathbf{Z}\alpha + \mathbf{u}$$

wo $\mathbf{y} = (y_i)_i$, $\mathbf{Z} = (Z_i)_{i,i}$, $\alpha = (\alpha_i)_i$, $\mathbf{u} = (u_i)_i$ mit $i=1, \dots, N$.

(vgl. dazu Anhang A1)

Welche stochastischen Bedingungen erfüllt nun das Störglied $\mathbf{u}$ aus (3.2)? Es ist

$$E\{\mathbf{u}\} = (E\{u_i\})_i = \mathbf{0}$$

$$E\{\mathbf{u}\mathbf{u}'\} = (E\{u_i u_k'\})_{i,k} = (\sigma_{ik}^2)_{i,k} = \Sigma \otimes \mathbf{I},$$

d.h. die Kovarianzmatrix ist nicht diagonal, wie sie sein müßte, wenn man die OLS-Methode korrekt anwenden wollte (vgl. (1bc)'in Abschnitt 1.2). Abschnitt 1.2. zeigt, daß im Fall von nichtdiagonaler Kovarianzmatrix die GLS-Methode die geeignete Methode ist. Anwendung der GLS-Methode (vgl. Definition 1.2) auf (3.2) ergibt folgende Schätzfunktion für α :

$$(3.3) \quad \hat{\alpha} = (\mathbf{Z}'(\Sigma^{-1} \otimes \mathbf{I})\mathbf{Z})^{-1} \mathbf{Z}'(\Sigma^{-1} \otimes \mathbf{I})\mathbf{y}$$

oder in etwas durchsichtigerer Schreibweise

1) In seinem Modell nennt Zellner die hier FI/OLS genannte Schätzmethode zweistufige Aitken-Schätzung.

$$(3.3') \quad \hat{\alpha} = (\sigma^{ih} z_i' z_h)^{-1}_{i,h} \left(\sum_{h=1}^N \sigma^{ih} z_i' y_h \right)_{i,1}$$

σ^{ih} ist dabei das (i,h) -te Element von Σ^{-1} .

Bemerkung 1: Hier wie auch beim 3SLS-Schätzverfahren tritt für die Praxis das Problem auf, daß die in den Schätzgleichungen auftretende Kovarianzmatrix Σ der kontemporären Störglieder nicht bekannt ist. Man behilft sich damit, daß man statt des wahren Σ eine Schätzung $\hat{\Sigma}$ in den Schätzgleichungen einsetzt. Zu einer konsistenten Schätzung für Σ kommt man, indem man das Modell mit irgendeiner konsistenten k -Klasse-Schätzung (meist 2SLS) vorschätzt und dann $\hat{\Sigma}$ als $\hat{\Sigma} = T^{-1} \hat{A}' Z' Z \hat{A} = T^{-1} \hat{U}' \hat{U}$ (natürlich unter Berücksichtigung der Nullrestriktionen) bestimmt. In Anlehnung an Dhrymes¹⁾ kann man dann von einer ausführbaren ('feasible') Schätzfunktion sprechen. Bei wahren (aber nicht bekanntem) Σ in der Schätzfunktion spricht man von einer nichtausführbaren ('non-feasible') Schätzfunktion. Um die Wahl der speziellen k -Klasse-Schätzfunktion, aus der $\hat{\Sigma}$ berechnet wird, frei zu lassen, werden hier die Full-Information-Schätzfunktionen als nichtausführbare Schätzfunktionen dargestellt. Durch das oben beschriebene Verfahren kommt man i.a. auch zu ausführbaren Schätzfunktionen.

(3.3) bzw. (3.3') bestimmen die FI/OLS (full information/ordinary least-squares/gewöhnliche Kleinst-Quadrate bei voller Information) -Schätzfunktion für α aus (3.2), also für alle Parameter des Modells (die nicht a-priori Null sind). In der Tat wird zur Schätzung von α die Information des gesamten Modells benutzt und, was wesentlich

1) vgl. Dhrymes, P.J.: Econometrics, a.a.O. S.152.

ist, in Form von Σ auch die stochastischen Zusammenhänge zwischen den verschiedenen Gleichungen. Warum man die Schätzfunktion (3.3) bzw. (3.3') mit Recht als Full-Information-Verallgemeinerung des OLS-Verfahrens bezeichnen kann, wird später (in Abschnitt 3.4.2. bzw. 4.) klar werden.

Definition 3.1: Die Schätzfunktion

$$\hat{\alpha} = \left(\sigma^{ih} z_i' z_h \right)_{i,h}^{-1} \left(\sum_h \sigma^{ih} z_i' y_h \right)_{i,1} = (Z' (\Sigma^{-1} \otimes I) Z)^{-1} Z' (\Sigma^{-1} \otimes I) y$$

heißt FI/OLS-Schätzfunktion für α aus (3.2).

In Analogie zu Satz 1.1 läßt sich folgender Satz formulieren:

Satz 3.1: Besteht Z in Gleichung (3.2) nur aus nichtstochastischen Größen, so ist die FI/OLS-Schätzfunktion aus Definition 3.1 ein BLUE.

Beweis: Folgt unmittelbar aus der Herleitung der FI/OLS-Schätzfunktion mit Hilfe der GLS-Methode und Satz 1.3.

Ebenso wie die Einzelgleichungs-OLS-Schätzfunktion ist jedoch auch die FI/OLS-Schätzfunktion im allgemeinen Fall nicht konsistent:

Satz 3.2: Falls im Modell (3.2) erklärende abhängige Variablen auftreten, ist die FI/OLS-Schätzfunktion (3.3) bzw. (3.3') inkonsistent.

Beweis: Erfolgt wie der Beweis von Satz 1.2, wobei der plim nun auf die in Gleichung (3.3') auftretenden Untermatrizen bzw. -vektoren anzuwenden ist. Der springende Punkt ist auch hier, daß plim $T^{-1} z_i' u_h \neq 0$. Auf eine genaue Herleitung kann verzichtet werden.

Bemerkung 2: Satz 3.1 zeigt, daß die OLS-Schätzfunktion nur dann ein BLUE ist (für die Parameter der i-ten Struk-

turgleichung), wenn man die i -te Strukturgleichung vollständig isoliert vom Restsystem betrachten kann, d.h. wenn man annehmen kann, daß keinerlei stochastische Zusammenhänge zwischen der i -ten und irgendeiner anderen Strukturgleichung bestehen.

3.2. Die 3SLS-Methode

Das 3SLS (three-stage least-squares/dreistufige Kleinst-Quadrate) -Schätzverfahren wurde von Zellner und Theil¹⁾ als Full-Information-Verallgemeinerung des 2SLS-Schätzverfahrens entwickelt, und zwar in Hinblick darauf, daß bei Verwendung der gesamten a-priori-Information des Modells ein "effizienteres" Verfahren als 2SLS möglich ist. In der Tat ist 3SLS im allgemeinen Fall effizienter als 2SLS in dem Sinne, daß die asymptotische 3SLS-Kovarianzmatrix kleiner als die asymptotische 2SLS-Kovarianzmatrix (des gesamten Systems) ist. In Abschnitt 3.5. wird darauf noch näher eingegangen. Ebenso wie das 2SLS- läßt sich auch das 3SLS-Verfahren auf verschiedene Weisen herleiten und motivieren.

3.2.1. 3SLS als IV-Schätzung²⁾

Zu schätzen ist das gesamte Modell (3.2). Wie bei 2SLS für das Z_i der i -ten Strukturgleichung, soll auch hier eine Hilfsvariablenmatrix $\mathbb{P}$ für das $\mathbb{Z}$ aus (3.2) gefunden werden. Für $\mathbb{P}$ muß dann gelten:

-
- 1) Zellner, A. und H. Theil: Three-Stage Least-Squares: Simultaneous Estimation of Simultaneous Equations, *Econometrica* 30, 1962, 54-78.
 - 2) 3SLS wurde zuerst von Madansky (Madansky, A.: On the Efficiency of Three-Stage Least-Squares Estimation, *Econometrica* 32, 1964, 51-56) als IV-Verfahren interpretiert.

$$(3.4) \quad \text{plim } T^{-1}P'u = 0 \quad \text{und}$$

$$(3.5) \quad \text{plim } T^{-1}P'Z = M \quad (|M| \neq 0),$$

damit die daraus gemäß (1.7) zu berechnende IV-Schätzfunktion konsistent ist. Als P wähle man in Analogie zu 2.2.1. und unter Berücksichtigung der Kovarianzmatrix der kon-temporären Störglieder die Matrix:

$$(3.6) \quad P = (\sigma^{ih} X(X'X)^{-1}X')_{i,h} (Z_1)_{i,1} = (\Sigma^{-1} \otimes X(X'X)^{-1}X')Z.$$

Daß die Bedingungen (3.4) und (3.5) mit dieser Wahl von P erfüllt sind, wird später noch im allgemeineren Fall der K -Matrix-Klasse-Schätzfunktion gezeigt. Ersetzt man Σ durch irgendeine konsistente k -Klasse-Schätzung $\hat{\Sigma}$ (vergl. Bemerkung 1 in 3.1.), so erkennt man leicht, daß das so aus (3.6) gewonnene $\hat{P}$ ebenfalls die Bedingungen (3.4) und (3.5) erfüllt.

Unter Beachtung von (1.6) erhält man mit dem P aus (3.6) folgende IV-Schätzfunktion für α :

$$(3.7) \quad \hat{\alpha} = (Z'(\Sigma^{-1} \otimes X(X'X)^{-1}X')Z)^{-1}Z'(\Sigma^{-1} \otimes X(X'X)^{-1}X')y \\ = (\sigma^{ih} Z_i'X(X'X)^{-1}X'Z_h)_{i,h}^{-1} \left(\sum_h \sigma^{ih} Z_i'X(X'X)^{-1}X'y_h \right)_{i,1}$$

und daraus folgende Definition:

Definition 3.2: Die Schätzfunktion (3.7) heißt (nichtausführbare 3SLS-Schätzfunktion für α aus (3.2). Als 'die' 3SLS-Schätzfunktion wird die ausführbare Schätzfunktion (3.7) bezeichnet, bei der Σ durch eine 2SLS-Vorschätzung bestimmt wird.

Aus der IV-Interpretation von 3SLS ergibt sich unmittelbar folgender Satz:

Satz 3.3: Die 3SLS-Schätzfunktion $\hat{\alpha}$ aus Definition 3.2 (ausführbar oder nichtausführbar) ist konsistent.

Bemerkung 1: Statt das oben beschriebene IV-Verfahren zu verwenden, könnte man die 3SLS-Schätzfunktion auch als GIV (generalized instrumental variable/verallgemeinertes Instrumenten-Variablen) -Verfahren interpretieren, wenn man als GIV (in Analogie zu GLS) die Schätzfunktion

$$\hat{\alpha} = (P'\Theta^{-1}Z)^{-1}P'\Theta^{-1}y$$

eingführt, sofern die Kovarianzmatrix Θ der Störglieder nicht diagonal ist. Im Falle des Gleichungssystems (3.2) ist dann $\Theta = \Sigma \otimes I$. Anstelle der Bedingungen (3.4) und (3.5) muß man nun die Bedingungen

$$\text{plim } T^{-1}P'\Theta^{-1}u = 0 \quad \text{und}$$

$$\text{plim } T^{-1}P'\Theta^{-1}Z = M \quad (|M| \neq 0)$$

nachweisen. Wahl von

$$P = X(X'X)^{-1}X'Z = \langle X(X'X)^{-1}X'Z_i \rangle_{i,i} = \langle (\hat{X}_i, X_i) \rangle_{i,i} \\ = \langle (\hat{Y}_i, X_i) \rangle_{i,i}$$

ergibt genau die 3SLS-Schätzfunktion (3.7) als GIV-Schätzfunktion.

Bemerkung 2: In der Bemerkung am Ende von 2.2.1. wurde gezeigt, daß 2SLS auch als OLS-Schätzfunktion des Ansatzes $y_i = P_i'\alpha_i + u_i^*$ betrachtet werden kann, wo P_i die Instrumentenmatrix aus der IV-Interpretation von 2SLS ist. In Analogie dazu läßt sich 3SLS als GLS-Schätzfunktion des Ansatzes $y = P\alpha + u^*$ betrachten mit $P = X(X'X)^{-1}X'Z$ aus der GIV-Interpretation von 3SLS und unter der Annahme $E\{u^*u^{*'}\} = \Sigma \otimes I$. Denn $P'(\Sigma^{-1} \otimes I)Z = P'(\Sigma^{-1} \otimes I)P$ wegen der Idempotenz von $X(X'X)^{-1}X'$. Diese GLS-Interpretation von 3SLS darf natürlich nicht mit der im folgenden Unterabschnitt behandelten verwechselt werden.

3.2.2. 3SLS als GLS-Schätzung

Aus derselben Motivation heraus, wie sie in 2.2.2. für 2SLS ausführlich dargelegt wurde, kann man einen transformierten Ansatz (3.2) GLS-schätzen und so zur 3SLS-Schätzfunktion gelangen. Man transformiert (in Analogie zu (2.4)) das Gleichungssystem (3.2) mit X' , indem man jeweils die entsprechenden Untervektoren bzw. -matrizen von y , Z und u mit X' prämultipliziert:

$$(3.8) \quad X'y = X'Z\alpha + X'u \quad , \quad \text{wo}$$

$$X'y = (X'y_1)_{1,1}, \quad X'Z = (X'Z_1)_{1,1} \quad \text{und} \quad X'u = (X'u_1)_{1,1} .$$

Als Kovarianzmatrix des Störgliedes in (3.8) bekommt man, wobei man wieder so tut, als sei X nichtstochastisch:

$$E\{X'u_1 u_1' X'\} = \left(E\{X'u_1 u_1' X'\} \right)_{1,k} = \left(\sigma_{ik} X'X \right)_{1,k} = \Sigma \otimes X'X .$$

Aus allgemeinen Regeln über das Kroneckerprodukt folgt

$$(\Sigma \otimes X'X)^{-1} = \Sigma^{-1} \otimes (X'X)^{-1} .$$

Mit $\Theta = \Sigma \otimes X'X$ ergibt also GLS-Schätzung des transformierten Ansatzes (3.8) (wobei $X'Z$ wie in (3.8) definiert ist):

$$\begin{aligned} \hat{\alpha} &= (Z'X(\Sigma^{-1} \otimes (X'X)^{-1})X'Z)^{-1} Z'X(\Sigma^{-1} \otimes (X'X)^{-1})X'y \\ &= \left(\sigma^{1h} z_1' X(X'X)^{-1} X' z_h \right)_{1,h}^{-1} \left(\sum_h \sigma^{1h} z_1' X(X'X)^{-1} X' y_h \right)_{1,1} \\ &= (Z'(\Sigma^{-1} \otimes X(X'X)^{-1})X'Z)^{-1} Z'(\Sigma^{-1} \otimes X(X'X)^{-1})X'y \end{aligned}$$

genau die (nichtausführbare) 3SLS-Schätzfunktion. Berechnen von $\hat{\alpha}$ über eine 2SLS-Vorschätzung ergibt 'die' 3SLS-Schätzfunktion.

Der Begriff "dreistufige" Kleinst-Quadrate-Schätzung ist mehr historisch bedingt: Man kann sich die Schätzung in drei Unterschätzungen unterteilt vorstellen: 1. OLS-Schätzung der reduzierten Form, 2. 2SLS-Schätzung von Σ , 3. IV-Schätzung mit der aus den beiden ersten Schätzungen gewonnenen IV-Matrix $\hat{P}$.

3.2.3. 3SLS als Ergebnis von Minimierungsprozessen

Als GLS-Schätzfunktion ist die 3SLS-Schätzfunktion natürlich auch Ergebnis eines Minimierungsprozesses. Wie bei 2SLS kann man dabei entweder die Spur oder die Determinante einer Matrix minimieren. 3SLS resultiert aus der GLS-Schätzung des transformierten Ansatzes (3.8). (3.8) GLS-schätzen bedeutet jedoch minimieren von

$$(3.9) \quad u'X\theta^{-1}X'u \quad \text{mit} \quad \theta^{-1} = \Sigma^{-1} \otimes (X'X)^{-1}$$

als Funktion von α .

(3.9) ist gleichbedeutend mit

$$(3.10) \quad u'(\Sigma^{-1} \otimes X(X'X)^{-1}X')u$$

(3.10) ist äquivalent zu

$$(3.11) \quad \sum_{i,h} \sigma^{ih} u_i' X(X'X)^{-1} X' u_h$$

(3.10) bzw. (3.11) entsprechen dem Minimanden (3.5) bei 2SLS. (3.11) ist aber nichts anderes als

$$\text{tr}(\Sigma^{-1} U' X(X'X)^{-1} X' U) \quad \text{oder}$$

$$(3.12) \quad \text{tr}(\Sigma^{-1} A' \tilde{Z}' \tilde{Z} A) \quad \text{mit} \quad \tilde{Z} = X(X'X)^{-1} X' Z$$

(Z, das in I, (2.1') definiert ist, darf nicht mit $\tilde{Z}$ verwechselt werden.)

Es ist nämlich mit $N^* = X(X'X)^{-1} X'$

$$\Sigma^{-1} U' N^* U = (\sigma^{ih})_{i,h} (u_h' N^* u_k)_{h,k} = (\sum_h \sigma^{ih} u_h' N^* u_k)_{i,k}$$

und die Spur davon ist nach Definition gleich

$$\sum_i \sum_h \sigma^{ih} u_h' N^* u_i = \sum_{i,h} \sigma^{ih} u_i' N^* u_h \quad \text{wegen der Symmetrie von } \Sigma^{-1}.$$

Dies ist genau (3.11).

(3.12) entspricht dem Minimanden (2.9) bei 2SLS und ist ebenfalls von Dhrymes¹⁾ angegeben.

Unter gewissen Einschränkungen läßt sich die 3SLS-Schätzfunktion auch durch Minimierung einer Determinante bestimmen, und zwar von

$$(3.13) \quad |U'X(X'X)^{-1}X'U|$$

wo $U = \left((y_1 - \tilde{z}_1' \alpha_1) \right)_{1,1}$, also unter Berücksichtigung aller

Nullrestriktionen.

Minimieren von (3.13) bedeutet offensichtlich minimieren von

$$(3.14) \quad \left| \left((y_1 - \tilde{z}_1' \alpha_1)' (y_h - \tilde{z}_h' \alpha_h) \right)_{1,h} \right|$$

mit $\tilde{z}_1 = X(X'X)^{-1}X'z_1$ (vgl. Dhrymes¹⁾ 2),

oder minimieren von

$$(3.15) \quad L := \frac{T}{2} \log |S| \quad \text{mit } S := (s_{1h}) = T^{-1} \left((y_1 - \tilde{z}_1' \alpha_1)' (y_h - \tilde{z}_h' \alpha_h) \right)_{1,h},$$

wegen der Monotonie von log.

Nun ist

$$\frac{\partial L}{\partial \alpha_j} = \frac{T}{2} \sum_{1,h} \frac{\partial \log |S|}{\partial s_{1h}} \frac{\partial s_{1h}}{\partial \alpha_j}$$

Da S symmetrisch ist, gilt nach bekannten Differentiationsregeln:

1) vgl. Dhrymes, P.J.: Small Sample and Asymptotic Relations..., a.a.O. S.363.

2) Die Modifikation bezüglich (3.13) macht man, weil für das

aus einer konsistenten Schätzung von α berechnete $\hat{S} = \frac{1}{T} \hat{U}'X(X'X)^{-1}X'\hat{U}$ wegen 1, 2(a),(b) und Folgerung 1 gilt: $\text{plim } \hat{S}^* = 0$.

$$\frac{\partial \log |S|}{\partial s_{ih}} = \begin{cases} s^{ih}, & \text{wenn } i=h \\ 2s^{ih}, & \text{wenn } i \neq h, \end{cases} \quad (S^{-1} = (s^{ih}))$$

so daß

$$\frac{\partial L}{\partial \alpha_j} = T \sum_{i,h} s^{ih} \frac{\partial s_{ih}}{\partial \alpha_j}.$$

Es ist aber

$$\frac{\partial s_{ih}}{\partial \alpha_j} = \begin{cases} 0 & \text{für } j \neq i \text{ und } j \neq h \\ -T^{-1} \tilde{z}_j' (y_h - \tilde{z}_h' \alpha_h) & \text{für } j=i, \end{cases} \quad \text{so daß}$$

$$\frac{\partial L}{\partial \alpha_j} = \sum_h s^{ih} \tilde{z}_i' (y_h - \tilde{z}_h' \alpha_h)$$

Nullsetzen von $\partial L / \partial \alpha_j$ ($j=1, \dots, N$) und umordnen ergibt:

$$(3.16) \quad (s^{ih} z_i' X(X'X)^{-1} X' z_h)_{i,h} \hat{\alpha} = (\sum_h s^{ih} z_i' X(X'X)^{-1} X' y_h)_{i,1}$$

Bis auf s^{ih} , das noch von α abhängt und damit (3.16) zu nichtlinearen Schätzgleichungen macht, sind dies die 3SLS-Schätzgleichungen. Berechnet man S , das in (3.15) definiert ist, aus einer Vorschätzung von α so, daß $\text{plim } \hat{S} = S$, oder ersetzt man s^{ih} in (3.16) durch $\hat{s}^{ih}$, so hat man die ausführbare oder die nichtausführbare 3SLS-Schätzfunktion.

Bemerkung: Auch die in 3.1. behandelte FI/OLS-Schätzfunktion läßt sich durch minimieren einer Determinante, und zwar der Determinante $|U'U|$ als Funktion der unbekannten Parameter, gewinnen, während man bei der Einzelgleichungs-OLS-Methode die Spur von $U'U$ minimiert. Zusammen mit (2.8) und (3.13) hat man nun folgendes schematisch zusammengefaßte interessante Ergebnis:

OLS = min tr(U'U)	FI/OLS = min U'U
2SLS = min tr(U'X(X'X) ⁻¹ X'U)	3SLS = min U'X(X'X) ⁻¹ X'U

Für (FI)/OLS gilt dabei $\text{plim } T^{-1}U'U = \Sigma$ (vgl. I, Folgerung 1), während man beim Minimanden von 3SLS die Modifikation (3.14) zu beachten hat. Anschaulich gesprochen berücksichtigen die Einzelgleichungsschätzverfahren, da die Spur die Summe der Diagonalelemente ist, nur die Varianzen, während die Systemverfahren mit der Determinante auch die Kovarianzen der Störglieder beachten. Mit der Transformationsmatrix $X(X'X)^{-1}X'$ bei 2SLS und 3SLS wird die Interdependenz berücksichtigt.

3.3. FISS-Methoden

Die den LISE-Verfahren entsprechenden Full-Information-Schätzverfahren könnte man als FISS (full-information simultaneous system)-Verfahren bezeichnen. Die den LIML- (LI/LGRV-, LVR-) Verfahren entsprechende FIML- (FI/LGRV-, LGVR-) Schätzverfahren führen wieder zu denselben Schätzgleichungen. Ebenso wie LISE hat auch FISS eine IV-Interpretation.

3.3.1. Die FIML-Methode¹⁾

Dem FIML (full-information maximum-likelihood) -Schätzverfahren liegt dieselbe (Log-)Likelihood-Funktion zugrunde wie LIML, nämlich die Funktion (2.11), die hier nochmals notiert sei:

$$(3.17) \quad L(A, \Sigma) = c + \log |\Gamma| - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \text{tr}(\Gamma' Z A \Sigma^{-1} A' Z')$$

1) erstmals dargestellt von Koopmans, T.C., Rubin, H. und R.B. Leipnik: Measuring the Equation Systems of Dynamic Economics, Cowles Commission Monograph 10, 1950, S.110-153.

vgl. auch Koopmans, T.C. und W.C. Hood: The Estimation of Simultaneous Linear Economic Relationships, Cowles Commission Monograph 14, 1953, S.143ff..

mit einer Konstanten c , $A = \begin{pmatrix} \Gamma \\ B \end{pmatrix}$, $Z = (Y, X)$.

Leitet man FIML ohne Annahme der Normalverteilung der Störglieder her, so heißt (3.17) Quasi-Maximum-Likelihood-Funktion. Da bei den Full-Information-Verfahren alle Parameter interessieren, muß (3.17) nicht mehr in dem Maße konzentriert werden, wie dies bei LIML der Fall war. Dies vereinfacht die Herleitung beträchtlich.

Es ist nun (3.17) zu maximieren. Dies geschieht wie üblich durch Differenzieren von $L(A, \Sigma)$ nach den unbekannten Parametern. Man kann auch hier schrittweise vorgehen, indem man L erst nach Σ und dann nach A differenziert¹⁾. Partielles Differenzieren von L nach Σ ergibt:

$$\frac{\partial L}{\partial \Sigma} = -\Sigma^{-1} + \Sigma^{-2} T^{-1} A' Z' Z A$$

nach bekannten Differentiationsregeln für Matrizen.

$$\text{Nullsetzen: } \frac{\partial L}{\partial \Sigma} = 0 \quad \Rightarrow$$

$$(3.18) \quad \Sigma(A) = T^{-1} A' Z' Z A \quad 2)$$

Einsetzen von (3.18) in (3.17) ergibt die konzentrierte Likelihood-Funktion

$$L^*(A) = c_1 + \log |\Gamma| - \frac{1}{2} \log |T^{-1} A' Z' Z A| \quad \text{oder}$$

$$(3.19) \quad L^*(A) = c_1 + \frac{1}{2} \log |\Gamma' \Gamma| - \frac{1}{2} \log |T^{-1} A' Z' Z A|,$$

wobei bisher noch keine Nullrestriktionen berücksichtigt wurden, da man annimmt, daß über Σ keine a-priori-Informationen außer der positiven Definitheit vorliegen. Aus Gründen, die später noch klar werden, maximiert man jedoch

1) Man kann zeigen, daß schrittweises Differenzieren und simultanes Differenzieren zu denselben Ergebnissen führt.

2) Man beachte die Analogie zu (2.21).

nicht (3.19), sondern den Ausdruck¹⁾

$$(3.20) \quad L = \frac{T}{2} \log |T^{-1} \Gamma' Y' \bar{M} Y \Gamma| - \frac{T}{2} \log |T^{-1} A' Z' Z A|$$

mit $\bar{M} = I - X(X'X)^{-1}X'$.

Dies ist gestattet, da $|T^{-1}Y'MY|$ bezüglich der Differentiation eine Konstante darstellt²⁾ und positiv definit ist. Außerdem sollen jetzt die Nullrestriktionen beachtet werden.

Mit

$$(3.21) \quad \begin{aligned} S &= (s_{ih}) = (T^{-1}(y_i - Z_i \alpha_i)'(y_h - Z_h \alpha_h))_{i,h} \quad (S^{-1} = (s^{ih})) \\ W &= (w_{ih}) = (T^{-1}(y_i - Y_i \delta_i)' \bar{M}(y_h - Y_h \delta_h))_{i,h} \quad (W^{-1} = (w^{ih})) \end{aligned}$$

erhält man (zu den Differentiationsregeln vergl. (3.15)):

$$\frac{\partial L}{\partial \alpha_j} = -T \sum_{i,h} s^{ih} \frac{\partial s_{ih}}{\partial \alpha_j} + T \sum_{i,h} w^{ih} \frac{\partial w_{ih}}{\partial \alpha_j}, \text{ wo}$$

$$\frac{\partial s_{ih}}{\partial \alpha_j} = \begin{cases} 0 & \text{für } j \neq i \text{ und } j \neq h \\ -T^{-1} Z_j' (y_h - Z_h \alpha_h) & \text{für } j = i \end{cases}$$

und

$$\frac{\partial w_{ih}}{\partial \alpha_j} = \begin{cases} 0 & \text{für } j \neq i \text{ und } j \neq h \\ -T^{-1} Z_j' \bar{M} (y_h - Z_h \alpha_h) & \text{für } j = i. \end{cases}$$

Dieser letzte Ausdruck kommt dadurch zustande, daß

$$X_j' \bar{M} (y_h - Z_h \alpha_h) = 0 \quad \text{ist.}$$

Zusammengefaßt erhält man also

$$\frac{\partial L}{\partial \alpha_j} = \sum_h s^{ih} Z_i' (y_h - Z_h \alpha_h) - \sum_h w^{ih} Z_i' \bar{M} (y_h - Z_h \alpha_h)$$

1) Man beachte die Analogie zu (2.22).

2) Chow, G.C.: Two Methods of Computing Full Information Maximum Likelihood Estimates in Simultaneous Stochastic Equations, International Economic Review 9, 1968, 100-112, erweitert mit $|T^{-1} \Gamma' Y' Y \Gamma|$.

Setzt man nun $\partial L / \partial \alpha_j = 0$ ($j=1, \dots, N$), ordnet um und verwendet Matrixschreibweise, so erhält man:

$$\left(Z_i' (s^{ih} I - w^{ih} \bar{M}) Z_h \right)_{i,h} \hat{\alpha} = \left(\sum_h Z_i' (s^{ih} I - w^{ih} \bar{M}) y_h \right)_{i,1}$$

oder

$$(3.22) \quad \left(s^{ih} Z_i' \left(I - \frac{w^{ih}}{s^{ih}} \bar{M} \right) Z_h \right)_{i,h} \hat{\alpha} = \left(\sum_h s^{ih} Z_i' \left(I - \frac{w^{ih}}{s^{ih}} \bar{M} \right) y_h \right)_{i,1}$$

(3.22) definiert die FIML-Schätzgleichungen für α aus (3.2). Natürlich ist das Gleichungssystem (3.22), da s^{ih} und w^{ih} (vgl. (3.21)) noch von α abhängen, hochgradig nicht-linear. Durch iterative Verfahren versucht man es in der Praxis zu lösen.

In Anlehnung an Abschnitt 3.2.3. soll hier die FIML-Schätzfunktion etwas anders definiert werden als allgemein in der Literatur üblich: Unter der FIML-Schätzfunktion wird i.a. die Schätzfunktion verstanden, für die die Funktion $L(A, \Sigma)$ aus (3.17) ihr absolutes Maximum erreicht, wobei A alle und Σ keine a-priori Restriktionen berücksichtigt. Wie man sieht, führt dies aber zu einem analytisch nicht lösbaren nichtlinearen Gleichungssystem. In der vorliegenden Arbeit sei nun durch (3.22) die nicht-ausführbare FIML-Schätzfunktion bestimmt (vgl. Bemerkung 1 in 3.1.), wenn man das s^{ih} , das in (3.22) direkt vor Z_i' steht, durch σ^{ih} ersetzt. Als ausführbare (oder linearisierte) FIML-Schätzfunktion wird die Schätzfunktion bezeichnet, die durch Auflösung von (3.22) nach $\hat{\alpha}$ entsteht, wenn s^{ih} und w^{ih} durch eine konsistente k -Klasse-Schätzung vorgeschätzt werden. Dieses Vorgehen ist dadurch gerechtfertigt, daß die so gewonnene Schätzfunktion die gleichen asymptotischen Eigenschaften¹⁾ wie die gewöhnlich als FIML bezeichnete Schätzfunktion hat.

Aus dem oben Gesagten ergibt sich folgende Definition:

1) vgl. dazu Dhrymes, P.J.: Small Sample and Asymptotic Relations. a.a.O.

Definition 3.3: Die Schätzfunktion

$$(3.23) \quad \hat{\alpha} = \left(\sigma^{ih} z_i' \left(I - \frac{w^{ih}}{s^{ih}} \bar{M} \right) z_h \right)^{-1}_{i,h} \left(\sum_h \sigma^{ih} z_i' \left(I - \frac{w^{ih}}{s^{ih}} \bar{M} \right) y_h \right)_{i,1}$$

($\bar{M} = I - X(X'X)^{-1}X'$, $w^{ih} = w^{ih}(\alpha)$ und $s^{ih} = s^{ih}(\alpha)$ sind in (3.21) definiert,) heißt (nichtausführbare) FIML-Schätzfunktion für α aus (3.2). Wird σ^{ih} , s^{ih} und w^{ih} durch eine konsistente Schätzung vorgeschätzt, so heißt sie ausführbare oder li-nearisierte FIML-Schätzfunktion.

Bemerkung 1: 'Die' FIML-Schätzfunktion ist nach Definition 3.3 die ausführbare FIML-Schätzfunktion, bei der σ^{ih} , s^{ih} , w^{ih} über das $\hat{\alpha}$ 'vor'geschätzt werden, das (3.20) absolut maximiert (falls ein solches $\hat{\alpha}$ existiert). Denn ein solches $\hat{\alpha}$ erfüllt die Schätzgleichungen (3.22).

Bemerkung 2: Die bezüglich Σ konzentrierte Likelihood-Funktion (3.20) ist offensichtlich genau das FI-Analogon zum Eingleichungsfall (2.22). Hier wird klar, warum man in (3.19) noch den Ausdruck $|T^{-1}Y'MY|$ einfügt. Man wäre also auch ohne die Gleichung (2.11) folgende mühsame Transformationstheorie zur konzentrierten Einzelgleichungs-Log-Likelihood-Funktion (2.22) gekommen, wenn man in (3.20) einfach A durch α_1 und Γ durch γ_1 ersetzt hätte. Allerdings wäre die Herleitung von LIML dann rein heuristisch gewesen. Nach (2.22) bzw. (3.20) trennen sich die Wege in der Herleitung von LIML und FIML: Bei LIML wird (2.22) weiter konzentriert durch partielles differenzieren nach β_1 , während man FIML aus Definition 3.3 durch sofortiges differenzieren nach dem gesamten A (also Γ und B) erhält. Durch das weitere konzentrieren bei LIML kommt man dann zu einer Schätzung für α_1 , die (2.22) tatsächlich maximiert. Man könnte aber nach Gleichung (2.22) auch ebenso

weiter vorgehen wie entsprechend nach Gleichung (3.20):
Sofortiges simultanes differenzieren von (2.22) bzw. (2.22')
nach α_1 ergibt¹⁾, wie man leicht nachweist:

$$\frac{\partial L^{**}(\alpha_1)}{\partial \alpha_1} = -s^{ii} Z_1' (y_1 - Z_1 \alpha_1) + w^{ii} Z_1' \bar{M} (y_1 - Z_1 \alpha_1) \quad ,$$

wo s^{ii} und w^{ii} wie in (3.21) definiert sind²⁾.

Nullsetzen ergibt:

$$(3.24) \quad Z_1' \left(I - \frac{w^{ii}}{s^{ii}} \bar{M} \right) Z_1 \hat{\alpha}_1 = Z_1' \left(I - \frac{w^{ii}}{s^{ii}} \bar{M} \right) y_1$$

$$\text{mit } \frac{w^{ii}}{s^{ii}} = \frac{s_{ii}}{w_{ii}} = \frac{(y_1 - Z_1 \alpha_1)' (y_1 - Z_1 \alpha_1)}{(y_1 - Z_1 \alpha_1)' \bar{M} (y_1 - Z_1 \alpha_1)} \quad .$$

Die Schätzgleichungen (3.24) sind also wieder nichtlinear.
In Analogie zu Definition 3.3 kann man jedoch folgende
neue Definition für LIML geben:

Definition 2.3': Die Schätzfunktion, die durch (3.24)
bestimmt wird, heißt (nichtausführbare) LIML-Schätzfunk-
tion für α_1 aus (2.1). Wird s^{ii} und w^{ii} durch eine kon-
sistente Schätzung vorgeschätzt, so heißt sie ausführbare
(oder linearisierte) LIML-Schätzfunktion.

'Die' LIML-Schätzfunktion (vgl. Definition 2.3) ist
dann unter der Normierung $\bar{\delta}_{ii} = -1$ die ausführbare LIML-
Schätzfunktion, die die Bedingung

$$\lambda = s_{ii}(\hat{\alpha}_1) / w_{ii}(\hat{\alpha}_1)$$

erfüllt mit dem λ , das die kleinste Wurzel des Eigenwert-
problems (2.28) ist.

-
- 1) Die Normierung $\bar{\delta}_{ii} = -1$ muß man jetzt natürlich beachten.
 - 2) wobei man jetzt allerdings W und S aus (3.21) diagonal
ansetzen kann, so daß $s^{ii} = 1/s_{ii}$ und $w^{ii} = 1/w_{ii}$.

Die durch Definition 2.3' erklärte LIML-Schätzfunktion ist das genaue Analogon zu der nach Definition 3.3 bestimmten FIML-Schätzfunktion. Sie ist eine k-Klasse-Schätzfunktion mit $\text{plim } k = \text{plim } (s_{ii}/w_{ii}) = \sigma_{ii}/\sigma_{ii} = 1$, also konsistent.

Umgekehrt könnte man natürlich auch die FI-Log-Likelihood-Funktion (3.20) weiter konzentrieren, indem man sie zunächst partiell bezüglich B maximiert und dann die konzentrierte Funktion bezüglich Γ maximiert. Man hätte dann das FI-Analogon zu Definition 2.3. Wie der nächste Abschnitt zeigen wird, stellen sich diesem Vorgehen jedoch unüberwindbare Schwierigkeiten entgegen.

3.3.2. Die LGVR-Methode

Wie 2.3.2 zeigte, kann LIML auch als LVR-Schätzmethode aufgefaßt werden. Man hatte dabei die Likelihood-Funktion (2.22) nochmals in Bezug auf β_1 zu konzentrieren. Dasselbe könnte man auch bei der (2.22) entsprechenden FI-Log-Likelihood-Funktion (3.20) versuchen. Partiell differenzieren von (3.20) nach (den Spalten von) B unter Beachtung der Nullrestriktionen (aber zunächst ohne die Normierung $\bar{y}_{ii} = -1$) ergibt:

$$\frac{\partial L}{\partial \beta_j} = -T \sum_{i,h} \bar{s}_{ih} \frac{\partial \bar{s}_{ih}}{\partial \beta_j}$$

wobei $\bar{s}_{ih}$ nun definiert ist durch $\bar{S} = (\bar{s}_{ih}) = (T^{-1} \bar{\alpha}_i' \bar{z}_i' \bar{z}_h' \bar{\alpha}_h)_{i,h}$

und

$$\frac{\partial \bar{s}_{ih}}{\partial \beta_j} = \begin{cases} 0 & \text{für } j \neq 1, j \neq h \\ -T^{-1} x_j' (\bar{y}_h \bar{y}_h + x_h \beta_h) & \text{für } j=1, \end{cases}$$

so daß nach Nullsetzen von $\partial L / \partial \beta_j$

$$\sum_{i,h} \bar{s}^{ih} x_i' x_h \beta_h = - \sum_{i,h} \bar{s}^{ih} x_i' \bar{y}_h \bar{\delta}_h$$

oder mit $\beta = (\beta_1', \dots, \beta_N')'$, $\bar{y} = (\bar{y}_1', \dots, \bar{y}_N')'$ in Matrixschreibweise:

$$(3.25) \quad \left(\bar{s}^{ih} x_i' x_h \right)_{i,h} \beta(\bar{y}) = - \left(\sum_h \bar{s}^{ih} x_i' \bar{y}_h \bar{\delta}_h \right)_{i,1}$$

(3.25) ist offensichtlich das FI-Analogon zu (2.23). Allerdings ist (3.25) ein nichtlineares Gleichungssystem, da $\bar{s}^{ih}$ noch von β abhängt. Man müßte schon dieses Gleichungssystem durch iterative (d.h. nicht-exakte) Methoden lösen. Im FI-Fall ist es also besser, (3.20) nicht weiter zu konzentrieren, sondern simultan nach A zu differenzieren, wie es in 3.3.1. geschehen ist.

Zu einem FI-Analogon des Varianzverhältnisses (2.26) bzw. (2.29) zu gelangen, aus dem die LVR-Methode erklärt wird, ist also exakt nicht möglich. Man kann jedoch bereits die (delogarithmierte) Likelihood-Funktion (3.20) als Varianzverhältnis auffassen und daraus, ohne Bezug auf das Likelihood-Prinzip zu nehmen, die LGVR¹⁾ (least generalized variance ratio/kleinstes verallgemeinertes Varianzverhältnis) -Schätzmethode herleiten, die natürlich zu denselben Schätzgleichungen wie FIML führt.

Delogarithmiert man (3.20), so erhält man

$$(3.26) \quad \frac{1}{s} = \frac{|T^{-1} \Gamma' Y' \bar{M} Y \Gamma|}{|T^{-1} (\Gamma' + X B)' (\Gamma' + X B)|} = \frac{|T^{-1} Y' \bar{M} Y|}{|T^{-1} \Gamma^{-1} (\Gamma' + X B)' (\Gamma' + X B) \Gamma^{-1}|}$$

Der Zähler rechts in (3.20) ist genau die Determinante der (OLS-) geschätzten Kovarianzmatrix der Störglieder der reduzierten Form, während der Nenner die bereits zu

1) nicht zu verwechseln mit LGRV

Beginn von 2.3.3. eingeführte verallgemeinerte Reststreuung der reduzierten Form ist. Maximieren von (3.26) bedeutet also minimieren des verallgemeinerten Varianzverhältnisses

$$(3.27) \quad s = \frac{|T^{-1}V'V|}{|T^{-1}\hat{V}'\hat{V}|} = \frac{|U'U|}{|r'V'Vr|}$$

als Funktion der Strukturparameter und unter Beachtung aller Nullrestriktionen. Die daraus resultierenden Schätzgleichungen sind die LGVR-Schätzgleichungen, die mit den FIML-Schätzgleichungen (3.23) übereinstimmen. LIML hingegen minimiert¹⁾

$$\frac{\text{tr}(U'U)}{\text{tr}(r'\hat{V}'\hat{V}r)}$$

als Funktion der unbekannten Parameter der i-ten Strukturgleichung (und unter Beachtung der Nullrestriktionen).

Das Schema aus der Bemerkung am Ende von 3.2.3. läßt sich jetzt wie folgt vervollständigen:

(3.28)

OLS = min tr(U'U)	FI/OLS = min U'U
2SLS = min tr(U'X(X'X) ⁻¹ X'U)	3SLS = min U'X(X'X) ⁻¹ X'U
LIML = min $\frac{\text{tr}(U'U)}{\text{tr}(r'\hat{V}'\hat{V}r)}$	FIML = min $\frac{ U'U }{ r'\hat{V}'\hat{V}r }$

3.3.3. Die FI/LGRV-Methode

Die FI/LGRV (full-information least generalized residual variance/kleinste verallgemeinerte Reststreuung bei voller Information) -Interpretation von FIML kann recht kurz abgehandelt werden: Sie ergibt sich aus delogarith-

1) dies folgt durch delogarithmieren von (2.22')

mieren von (3.19) (bzw. daraus, daß man in (3.27) nur den Zähler betrachtet): Minimieren der verallgemeinerten Reststreuung¹⁾ $|T^{-1}V'V|$ als Funktion der Strukturparameter unter Beachtung der Nullrestriktionen ergibt die FI/LGRV-Schätzgleichungen, die natürlich identisch mit den FIML-Schätzgleichungen sind. Das FI/LGRV-Verfahren ist nicht das genaue FI-Pendant zu LI/LGRV, denn dort (vgl. 2.3.3.) minimiert man ja durch differenzieren nach den reduzierten-Form-Parametern unter Nebenbedingungen, die die eindeutige Beziehung zwischen reduzierter Form und Strukturform herstellen. Auf die genaue FI-Verallgemeinerung dieses Vorgehens wird hier verzichtet²⁾.

3.3.4. FISS als IV-Schätzung

Aus Definition 3.3 wird die IV-Interpretation von FIML unmittelbar klar³⁾: $\hat{\alpha}$ aus (3.23) ist eine IV-Schätzfunktion mit der Instrumentenmatrix

$$(3.29) \quad P = \left(\sigma^{ih} \left(I - \frac{w^{ih}}{s^{ih}} \bar{M} \right) \right)_{i,h} \left(z_1 \right)_{i,1}.$$

P hängt dabei noch von den unbekannten Parametern ab, da s^{ih} und w^{ih} Funktionen von α sind (vgl. (3.21)). Im ausführbaren Fall werden s^{ih} , w^{ih} und σ^{ih} durch eine konsistente Schätzung vorgeschätzt.

Sind nun mit dem P aus (3.29) sowohl im ausführbaren als auch im nichtausführbaren Fall die Bedingungen

-
- 1) Man vergleiche hierzu die von Brown, T.M.: Simultaneous Least Squares: A Distribution Free Method of Equation System Structure Estimation, International Economic Review 1, 1960, 173-191, eingeführte Simultane Kleinst-Quadrate-Methode, die $\text{tr}(V'V)$ als Funktion der Strukturparameter minimiert.
 - 2) vgl. dazu auch Schönfeld, P.: Methoden der Ökonometrie, Bd.II, a.a.O. S.283.
 - 3) vgl. dagegen die von Hausmann, J.A.: An Instrumental Variable Approach to Full Information Estimators for Linear and Certain Nonlinear Econometric Models, Econometrica 43, 1975, 727-738, angegebene IV-Interpretation von FIML.

$$\text{plim } T^{-1}P'u = 0 \quad \text{und}$$

$$\text{plim } T^{-1}P'Z = M \quad (|M| \neq 0)$$

erfüllt? Man könnte diese Frage auf die entsprechende Frage für 3SLS zurückführen, wenn $\text{plim } (w^{ih}/s^{ih}) = 1$ wäre (vgl. (3.6)). Dann wären nämlich asymptotisch die Instrumentenmatrizen von 3SLS und FIML gleich. Es ist in der Tat

$$\text{plim } S^{-1} = \text{plim } (T^{-1}U'U)^{-1} = \Sigma^{-1} \quad \text{und}$$

$$\text{plim } W^{-1} = (\Gamma' \text{plim } T^{-1} \hat{V} \hat{V}' \Gamma)^{-1} = (\Gamma' \Gamma^{-1} \Sigma^{-1} \Gamma^{-1} \Gamma)^{-1} = \Sigma^{-1}$$

so daß $\text{plim } s^{ih} = \text{plim } w^{ih}$ und damit $\text{plim } (w^{ih}/s^{ih}) = 1$. Aus der IV-Interpretation von FIML ergibt sich also folgender

Satz 3.4: Die FIML-Schätzfunktion $\hat{\alpha}$ aus Definition 3.3 (ausführbar oder nichtausführbar) ist konsistent¹⁾.

Bemerkung: Ebenso wie 3SLS (vgl. Bemerkung 1 in 3.2.1.) könnte man auch FIML als GIV-Schätzung betrachten mit der Instrumentenmatrix

$$P = \left(I - \frac{w^{ih}}{s^{ih}} \bar{M} \right)_{i,h} \begin{pmatrix} z_i \\ 1 \end{pmatrix}_{i,i},$$

so daß die FIML-Schätzfunktion sich schreiben ließe als

$$(3.30) \quad \hat{\alpha} = (P'(\Sigma^{-1} \otimes I)Z)^{-1} P'(\Sigma^{-1} \otimes I)y$$

mit obigem P .

1) Die Konsistenz 'der' FIML-Schätzfunktion ist damit natürlich nicht gezeigt. Denn die Konsistenz 'der' FIML-Schätzfunktion wird schon zur Berechnung von $\hat{s}^{ih}$, $\hat{s}^{ih}$, $\hat{w}^{ih}$ benötigt (vgl. Bemerkung 1 in 3.3.1.). Man muß die Konsistenz in diesem Fall auf andere Weise nachweisen. Dies ist in der Tat möglich (vgl. Sargan, J.D.: Three-Stage Least Squares and Full Maximum Likelihood Estimates, Econometrica 1964, 77-81.)

3.4. Die K-Matrix-Klasse

FI-Verallgemeinerungen der k-Klasse wurden von Roy und Srivastava¹⁾ (Generalized double-k-class/GDkC) und von Savin²⁾ (System's k-class/SkC) vorgeschlagen. Sie enthalten jedoch nicht die FIML-Schätzfunktion als Spezialfall. Die hier behandelte K-Matrix-Klasse (KMC)³⁾ erfüllt diese Bedingung. Außerdem zeigt sich, daß GDkC und SkC Unterklassen der später noch zu behandelnden Doppel-K-Matrix (DKMC) sind. Nach Definition und IV-Interpretation von KMC im 1. Unterabschnitt wird gezeigt, daß alle FI-Verfahren zur (symmetrischen) K-Matrix-Klasse gehören. Als FI-Verallgemeinerungen der h- und Doppel-k-Klasse werden h-Vektor- und Doppel-K-Matrix-Klasse eingeführt.

3.4.1. Herleitung und Definition der K-Matrix-Klasse (IV-Interpretation von KMC)

Da sowohl 3SLS als auch FIML zur K-Matrix-Klasse (K-matrix-class/KMC) gehören sollen, legen Definition 3.2 und Definition 3.3 folgende Definition nahe:

Definition 3.4: Die N^2 -parametrische Schar von Schätzfunktionen für α aus Gleichung (3.2), die durch die Schätzgleichungen

$$(3.31) \quad \left(\sigma^{ih} z_i' (I - k_{ih} \bar{M}) z_h \right)_{i,h} \hat{\alpha}(K) = \left(\sum_h \sigma^{ih} z_i' (I - k_{ih} \bar{M}) y_h \right)_{i,1}$$

1) Roy, A.R. und V.K. Srivastava: Generalized Double k-Class Estimators, Journal of the Indian Statistical Association 4, 1966, 38-46.

2) Savin, N.E.: Systems k-Class Estimators, Econometrica 41, 1973, 1125-1136.

3) vgl. Scharf, W.: K-Matrix-Class Estimators..., a.a.O. 41-50.

($\bar{M} = I - X(X'X)^{-1}X'$) bestimmt ist, heißt (nichtausführbare) K-Matrix-Klasse. Dabei ist $K = (k_{ih})$ eine beliebige stochastische oder nichtstochastische $N \times N$ -Matrix, die sogar von den unbekannten Parametern α abhängen kann. Wird σ^{ih} und (falls nötig) K durch eine konsistente Vorschätzung berechnet, so definiert (3.31) die ausführbare oder linearisierte K-Matrix-Klasse.

Wie die k-Klasse kann auch die K-Matrix-Klasse in der suggestiveren Form

$$(3.32) \quad \left(\sigma^{ih} \begin{pmatrix} Y_i' Y_h - k_{ih} \hat{v}_i' \hat{v}_h & Y_i' X_h \\ X_i' Y_h & X_i' X_h \end{pmatrix} \right)_{i,h} \hat{\alpha} = \left(\frac{\sum_h \sigma^{ih} (Y_i' - k_{ih} \hat{v}_i') Y_h}{\sum_h \sigma^{ih} X_i' Y_h} \right)_{i,i}$$

geschrieben werden mit $\hat{v}_i = Y_i - \hat{Y}_i = (I - X(X'X)^{-1}X')Y_i$, dem OLS-geschätzten Residuenanteil der reduzierten Form, der zu Y_i gehört.

Auch die Schätzungen der K-Matrix-Klasse besitzen eine naheliegende IV-Interpretation¹⁾ mit der Instrumentenmatrix

$$(3.33) \quad P(K) = \left(\sigma^{ih} (I - k_{ih} \bar{M}) \right)_{i,h} \begin{pmatrix} Z_i \end{pmatrix}_{i,i}.$$

Die Bedingungen

$$\text{plim } T^{-1} P'(K) u = 0 \quad \text{und}$$

$$\text{plim } T^{-1} P'(K) Z = M \quad (|M| \neq 0)$$

1) Als GIV-Schätzung (vgl. Bem. 1 in 3.2.1.) würde die Instrumentenmatrix folgendermaßen aussehen:

$$P(K) = \left((I - k_{ih} \bar{M}) \right)_{i,h} \begin{pmatrix} Z_i \end{pmatrix}_{i,i}$$

und (3.31) ließe sich schreiben als

$$P'(K) (Z^{-1} \otimes I) Z \hat{\alpha} = P'(K) (Z^{-1} \otimes I) y.$$

(und damit die Konsistenz der KMC-Schätzungen) sind jedoch nicht für jedes K erfüllt. Entsprechend Satz 2.4 für die k -Klasse gibt folgender Satz für KMC Auskunft darüber, für welche Matrizen K obige Bedingungen erfüllt sind, die KMC-Schätzungen also tatsächlich IV-Schätzungen darstellen.

Satz 3.5: Sei die Menge der in Definition 3.4 zugelassenen $N \times N$ -Matrizen K .

$\mathcal{M}_1$ sei die Untermenge von $\mathcal{M}$, für deren Elemente gilt: $\text{plim}(K-E)=0$ (E ist die Matrix die nur aus Einsen besteht).

$\mathcal{M}_2$ sei die Untermenge von $\mathcal{M}$, deren Elemente durch die Schätzgleichungen (3.31) eine konsistente Schätzfunktion bestimmen.

Dann gilt

$$\mathcal{M}_1 = \mathcal{M}_2$$

Beweis: Der Beweis erfolgt entsprechend dem von Satz 2.4.

1. Sei $K \in \mathcal{M}_1$. Dann gilt mit $P(K) = (P_{ih}(k_{ih}))_{i,h}$

$$\begin{aligned} \text{plim } T^{-1}P'(K')Z &= (\text{plim } T^{-1}P'_{ih}(k_{ih})Z_h)_{i,h} = \\ (3.34) \quad &= \left(\sigma_{ih} \begin{pmatrix} \Pi'_i M_h + \text{plim}(1-k_{ih})D'_i \Sigma D_h, \Pi'_i M_h \\ M'_i \Pi_h, M_{ih} \end{pmatrix} \right)_{i,h} \end{aligned}$$

Dies erhält man durch ersetzen von Y_i durch $XW_i + V_i$ und Anwendung des plim s unter Beachtung von I, Annahme 2(a) und Folgerung 1.

$$(3.35) \quad \text{plim } T^{-1}P'(K')u = \left(\sum_h \sigma_{ih} \text{plim}(1-k_{ih})D'_i \sigma_h \right)_{i,1}$$

Wegen $\text{plim}(K'-E) = (\text{plim}(k_{ih}-1))_{i,h} = 0$ ist (3.35) gleich Null und (3.34) eine nichtsinguläre Matrix. Die beiden

Bedingungen für eine IV-Schätzung sind also für jedes $K \in \mathcal{M}_1$ erfüllt, d.h. eine K-Matrix-Klasse-Schätzung mit $K \in \mathcal{M}_1$ ist konsistent bzw. $K \in \mathcal{M}_2$.

2. Sei umgekehrt $K \in \mathcal{M}_2$. Dann ist $\hat{\alpha}(K^*)$ konsistent und es gilt

$$(3.36) \quad \text{plim } \hat{\alpha}(K^*) = \alpha + (\sigma^{ih} \text{plim } T^{-1} P'_{ih} (k_{ih}^*) z_h)^{-1}_{i,h} \\ (\sum_h \sigma^{ih} \text{plim } T^{-1} P'_{ih} (k_{ih}^*) u_h)_{i,1} \\ = \alpha, \text{ wegen der Konsistenz.}$$

Der zweite Summand auf der rechten Seite von (3.36) muß also Null sein. Dies ist genau dann der Fall, wenn $\text{plim } T^{-1} P'_{ih} (k_{ih}^*) z_h$ existiert und $\text{plim } T^{-1} P'_{ih} (k_{ih}^*) u_h = 0$ ($i, h=1, \dots, N$). Wie (3.34) und (3.35) zeigen, ist eine notwendige Bedingung hierfür, daß $\text{plim}(k_{ih}^* - 1) = 0$ ($i, h=1, \dots, N$), d.h. $K \in \mathcal{M}_1$,
q.e.d.

3.4.2. FI/OLS, 3SLS und FISS als K-Matrix-Klasse-Schätzungen

Daß alle FI-Schätzverfahren zur K-Matrix-Klasse gehören, zeigt man nach dem Vorangegangenen leicht, wenn man die Definitionen 3.1, 3.2, 3.3 und 3.4 vergleicht. Zusammenfassend kann man folgenden Satz formulieren:

Satz 3.6: Aus den KMC-Schätzgleichungen (3.31) erhält man bei verschiedener Wahl der $N \times N$ -Matrix K die folgenden (ausführbaren oder nichtausführbaren) Schätzfunktionen für α aus Gleichung (3.2): Für

a) $K = 0$ die FI/OLS-Schätzfunktion¹⁾

1) Hier wird u.a. der Begriff FI/OLS deutlich, denn die Einzelgleichungs-OLS-Methode erhält man analog aus der k-Klasse durch Nullsetzen von k .

b) $K = E$ die 3SLS-Schätzfunktion (E ist die Matrix, die nur aus Einsen besteht)

c) $K = \begin{pmatrix} w^{ih} \\ s^{ih} \end{pmatrix}$ die FISS (FIMT-, FI/LGRV-, LGVR-) -Schätzfunktion. Dabei ist

$$s^{-1} = (s^{ih}) = (T^{-1}(y_i - z_i \alpha_i)' (y_h - z_h \alpha_h))_{i,h}^{-1}$$

$$w^{-1} = (w^{ih}) = (T^{-1}(y_i - y_i \delta_i)' \bar{M}(y_h - y_h \delta_h))_{i,h}^{-1}$$

d) Im genau identifizierten Fall des gesamten Modells erhält man außerdem für

$K = E$ die ILS-Schätzfunktion (für das gesamte System)

d) erhält man, da im genau identifizierten Fall alle Schätzfunktionen mit ILS zusammenfallen, was noch in Abschnitt 5.2 gezeigt wird.

Bemerkung: Da alle Matrizen in Satz 3.6 symmetrisch sind, gehören die FI-Schätzfunktionen bereits zu der $N(N+1)/2$ -parametrischen Schar von Schätzfunktionen, die durch Definition 3.4 mit symmetrischem K bestimmt ist. Die entsprechende Unterklasse von KMC könnte man als symmetrische K -Matrix-Klasse bezeichnen.

3.4.3. h-Vektor- und Doppel-K-Matrix-Klasse

Als FI-Äquivalent zu h - und Doppel- k -Klasse sollen hier h -Vektor-Klasse (hVC) und Doppel- K -Matrix-Klasse ($DKMC$) eingeführt werden. Die Herleitung der h -Vektor-Klasse erfolgt analog zu der der h -Klasse. Man betrachte wieder das Gleichungssystem (3.2) und bereinige dort alle Y_i ($i=1, \dots, N$) um einen (bei jedem Y_i möglicherweise verschiedenen) Anteil der für Y_i relevanten OLS-geschätzten reduzierten-Form-Residuen $\hat{v}_i$, so daß man aus (3.2) folgenden Ausdruck bekommt:

$$(3.37) \quad y = (Y-h\hat{V})_1 + \alpha\beta + u^* = \begin{pmatrix} Y-h\hat{V} \\ \alpha \end{pmatrix} \alpha^* + u^*$$

mit $y = (y_1)_{1,1}$, $Y-h\hat{V} = (Y_1-h_1\hat{V}_1)_{1,1}$, $\alpha = (\alpha_1)_{1,1}$,

$$\gamma = (\gamma_1)_{1,1}, \quad \beta = (\beta_1)_{1,1}, \quad u^* = (u_1)_{1,1}, \quad \alpha^* = \begin{pmatrix} \gamma \\ \beta \end{pmatrix}.$$

GLS-Schätzung von (3.37) unter der Annahme $E\{u^*u^{*\prime}\} = \Sigma \otimes I$ ergibt:

$$(3.38) \quad \begin{pmatrix} (Y-h\hat{V})' \\ \alpha' \end{pmatrix} (\Sigma^{-1} \otimes I) (Y-h\hat{V}, \alpha) = \begin{pmatrix} (Y-h\hat{V})' \\ \alpha' \end{pmatrix} (\Sigma^{-1} \otimes I) y.$$

Multipliziert man obige Schätzgleichung (3.38) aus, so erhält man für die einzelnen Untermatrizen:

$$(Y-h\hat{V})' (\Sigma^{-1} \otimes I) (Y-h\hat{V}) = (\sigma^{1k} (Y_1-h_1\hat{V}_1)' (Y_k-h_k\hat{V}_k))_{1,k} =$$

$$= (\sigma^{1k} (Y_1'Y_k - (h_1+h_k-h_1h_k)\hat{V}_1'\hat{V}_k))_{1,k}$$

$$\alpha' (\Sigma^{-1} \otimes I) (Y-h\hat{V}) = (\sigma^{1k} \alpha_1' (Y_k-h_k\hat{V}_k))_{1,k} = (\sigma^{1k} \alpha_1' Y_k)_{1,k}$$

$$\alpha' (\Sigma^{-1} \otimes I) \alpha = (\sigma^{1k} \alpha_1' \alpha_k)_{1,k}$$

$$(Y-h\hat{V})' (\Sigma^{-1} \otimes I) y = (\sum_k \sigma^{1k} (Y_1-h_1\hat{V}_1)' y_k)_{1,1}$$

$$\alpha' (\Sigma^{-1} \otimes I) y = (\sum_k \sigma^{1k} \alpha_1' y_k)_{1,1}$$

Insgesamt lassen sich die Schätzgleichungen also schreiben als

$$\begin{pmatrix} (\sigma^{1k} (Y_1'Y_k - (h_1+h_k-h_1h_k)\hat{V}_1'\hat{V}_k))_{1,k} & , & (\sigma^{1k} \alpha_1' \alpha_k)_{1,k} \\ (\sigma^{1k} \alpha_1' Y_k)_{1,k} & , & (\sigma^{1k} \alpha_1' y_k)_{1,k} \end{pmatrix} \hat{\alpha}^* =$$

$$(3.39)$$

$$(3.39) \quad = \begin{pmatrix} \left(\sum_k \sigma^{ik} (Y_i - h_i \hat{v}_i)' y_k \right)_{i,1} \\ \left(\sum_k \sigma^{ik} X_i' y_k \right)_{i,1} \end{pmatrix}$$

oder bei anderer Zusammenfassung der Parameter:

$$(3.39') \quad \left(\sigma^{ik} \begin{pmatrix} Y_i' y_k - (h_i + h_k - h_i h_k) \hat{v}_i' \hat{v}_k & Y_i' X_k \\ X_i' y_k & X_i' X_k \end{pmatrix} \right)_{i,k} \alpha =$$

$$= \begin{pmatrix} \sum_k \sigma^{ik} \begin{pmatrix} Y_i' - h_i \hat{v}_i' \\ X_i' \end{pmatrix} y_k \\ \sum_k \sigma^{ik} X_i' X_k \end{pmatrix}_{i,1}$$

mit $\alpha = (\gamma_1', \beta_1', \dots, \gamma_N', \beta_N')'$.

Macht man in (3.39) bzw. (3.39') die Substitution

$$h \rightarrow 1-h^* \quad (\text{d.h. } h_i \rightarrow 1-h_i^* \quad (i=1, \dots, N)),$$

so erhält man die Schätzgleichungen in "symmetrischer" Form: Auf der linken Seite von (3.39') ergibt sich dann jeweils

$$h_i + h_k - h_i h_k = (1-h_i^*) + (1-h_k^*) - (1-h_i^*)(1-h_k^*) =$$

$$= 2-h_i^* - h_k^* - 1 + h_i^* + h_k^* - h_i^* h_k^* = 1-h_i^* h_k^*$$

und auf der rechten Seite von (3.39')

$$h_i = 1-h_i^*.$$

Die Schätzgleichungen (3.39') definieren die h -Vektor-Klasse, die durch Variation des $N \times 1$ -Vektors h eine N -parametrische Schar von Schätzfunktionen bestimmt.

Die Doppel-K-Matrix-Klasse (DKMC) ist die Verallgemeinerung der K-Matrix-Klasse, die in Definition 3.4 statt nur einer Matrix K links und rechts von (3.31) verschiedene (voneinander unabhängige) Matrizen K_l und K_r zuläßt. Zusammengefaßt, und wenn man in (3.39') wieder beachtet, daß $\hat{v}_i = (I - X(X'X)^{-1}X')Y_i = \tilde{M}Y_i$, erhält man die folgende Definition:

Definition 3.5: a) Die N-parametrische Schar von Schätzfunktionen für α aus (3.2), die durch die Schätzgleichungen

$$(3.40) \left(\sigma^{ik} z_i' (I - (1-h_1 h_k) \bar{M}) z_k \right)_{i,k} \hat{\alpha}(h) = \left(\sum_k \sigma^{ik} z_i' (I - (1-h_1) \bar{M}) y_k \right)_{i,1}$$

bestimmt ist, heißt h-Vektor-Klasse (hVC).

b) Die $2N^2$ -parametrische Schar von Schätzfunktionen für α aus (3.2), die durch die Schätzgleichungen

$$(3.41) \left(\sigma^{ih} z_i' (I - k_{1,ih} \bar{M}) z_h \right)_{i,h} \hat{\alpha}(K_1, K_R) = \left(\sum_h \sigma^{ih} z_i' (I - k_{r,ih} \bar{M}) y_h \right)_{i,1}$$

bestimmt ist, heißt Doppel-K-Matrix-Klasse (DKMC)¹⁾.

$h = (h_i)_{i,1}$ in a) ist dabei ein beliebiger (stochastischer oder nichtstochastischer) $N \times 1$ -Vektor, $K_1 = (k_{1,ih})_{i,h}$ und $K_R = (k_{r,ih})_{i,h}$ in b) sind beliebige (stochastische oder nichtstochastische) voneinander unabhängige $N \times N$ -Matrizen.

KMC ist offensichtlich die Unterklasse von DKMC²⁾, für die $K_1 = K_R$. DKMC enthält damit natürlich alle Schätzfunktionen, die auch KMC enthält. Die h-Vektor-Klasse ist ebenfalls eine Unterklasse von DKMC (mit der speziellen Wahl von $K_1 = (1-h_1 h_k)_{i,k}$ und $K_R = ((1-h_1)_{i,1}, \dots, N\text{-mal} \dots, (1-h_1)_{i,1})$, nicht jedoch von KMC³⁾).

Welche Schätzfunktionen haben KMC und hVC gemeinsam?

-
- 1) Unter der symmetrischen Doppel-K-Matrix-Klasse versteht man dann die Unterklasse von DKMC mit symmetrischen K_1 und K_R .
 - 2) Roy und Srivastavas (Roy, A.R. und V.K. Srivastava: a.a.O. 38-46) GDkC (generalized double k-class) ist die Unterklasse von DKMC, für die K_1 identische Spalten hat, Savins (Savin, N.E.: a.a.O. 1125-1136) systems k-class ist die Unterklasse von KMC, für die K aus identischen Elementen besteht.
 - 3) Damit ist hVC auch eine Unterklasse von GDkC.

Dies sind offensichtlich (vgl. (3.40)) genau die Schätzfunktionen, für die simultan erfüllt ist:

$$\begin{aligned} 1 - h_i h_k &= 1 - h_i & (i, k=1, \dots, N) \text{ oder} \\ (3.42) \quad h_i (1 - h_k) &= 0 & (i, k=1, \dots, N). \end{aligned}$$

Die beiden Lösungen von (3.42) sind $h = 1$ und $h = 0$. Einsetzen dieser Werte in (3.40) ergibt die FI/OLS- bzw. die 3SLS-Schätzgleichungen. Nur diese beiden Schätzfunktionen sind also sowohl in KMC als auch in hVC enthalten.

Lemma: Aus den h-Vektor-Klasse-Schätzgleichungen (3.40) erhält man bei verschiedener Wahl des $N \times 1$ -Vektors h die folgenden Schätzfunktionen für α aus (3.2): Für

- $h = 1$ die FI/OLS-Schätzfunktion
- $h = 0$ die 3SLS-Schätzfunktion.

3.5. Eigenschaften der Schätzungen der K-Matrix-Klasse

In Satz 3.5 wurde bereits für eine Unterklasse von KMC die Konsistenz nachgewiesen (nämlich für die Unterklasse von KMC, für die $\text{plim}(K-E)=0$). Für eine noch kleinere Unterklasse von KMC wird nun ohne Beweis¹⁾ die asymptotische Verteilung (zu diesem Begriff vgl. Abschnitt 2.5) angegeben. Aus der Kovarianzmatrix dieser asymptotischen Verteilung kann man dann schließen, daß die FI-Schätzverfahren, die dieser Unterklasse von KMC angehören, asymptotisch "relativ effizient" bezüglich der - der entsprechenden Unterklasse der k-Klasse angehörenden - LI-Schätzverfahren sind.

1) zum Beweis, der auf den entsprechenden Beweis für 3SLS (vgl. Schönfeld, P.: Methoden der Ökonometrie, Bd.II, a.a.O.S.251) zurückgeführt wird, siehe Scharf, W.: K-Matrix-Class Estimators..., a.a.O.S.48ff.

Analog zu Satz 2.6 erhält man folgenden Satz über die asymptotische Verteilung der KMC-Schätzfunktionen:

Satz 3.7: $\mathcal{M}$ sei die Menge der in Definition 3.4 zugelassenen $N \times N$ -Matrizen K .

$\mathcal{M}'$ sei die Untermenge von $\mathcal{M}$, für deren Elemente gilt:

$\text{plim} \sqrt{T}(K-E) = 0$ (E ist die Matrix, die nur aus Einsen besteht).

$\mathcal{M}''$ sei die Untermenge von $\mathcal{M}$, deren Elemente durch die Schätzgleichungen (3.31) eine Schätzfunktion mit der

asymptotischen Normalverteilung $N(\alpha, T^{-1}(\sigma_{ih} \bar{Q}_{ih})^{-1}_{i,h})$

bestimmen. $\bar{Q}_{ih}$ ist dabei die (nichtstochastische) Matrix

$\text{plim} T^{-1} Z_1' X (X'X)^{-1} X' Z_h$.

Dann gilt

$\mathcal{M}' \subset \mathcal{M}''$ ($\mathcal{M}'$ ist Untermenge von $\mathcal{M}''$).

Wegen $\text{plim} \sqrt{T}(E-E) = 0$ hat die 3SLS-Schätzfunktion die asymptotische Kovarianzmatrix aus Satz 3.7. Es läßt sich nachweisen¹⁾, daß auch die FIML-Schätzfunktion diese asymptotische Kovarianzmatrix besitzt. 3SLS und FIML sind also asymptotisch äquivalent.

Um die asymptotische Kovarianzmatrix der k -Klasse-Schätzfunktionen mit der der KMC-Schätzfunktionen vergleichen zu können, muß man die k -Klasse für das gesamte System betrachten. Diese ist entsprechend (2.39) definiert durch die Schätzgleichungen

$$(3.43) \quad (Z_1' (I - k_i \bar{M}) Z_1)_{i,i}^{-1} \hat{\alpha} = (Z_1' (I - k_i \bar{M}) y_1)_{i,i}^{-1}$$

Ihre asymptotische Kovarianzmatrix lautet, falls

$\text{plim} \sqrt{T}(k_i - 1) = 0$ ($i=1, \dots, N$):

1) vgl. Sargan, J.D.: Three-Stage Least Squares and Full Maximum Likelihood Estimates, a.a.O. 77-81.

$$(3.44) \quad Q_2 = (\bar{Q}_{11}^{-1})_{i,i} (\sigma_{ih} \bar{Q}_{ih})_{i,h} (\bar{Q}_{11}^{-1})_{i,i},$$

wo $\bar{Q}_{ih}$ in Satz 3.7 erklärt ist.

Satz 3.8: K_1 sei die Unterklasse von KMC, für die gilt:
 $\text{plim} \sqrt{T}(K-E) = 0$, und deren Schätzfunktionen damit die
 asymptotische Kovarianzmatrix $Q_1 = (\sigma_{ih} \bar{Q}_{ih})_{i,h}^{-1}$ aus
 Satz 3.7 besitzen.

K_2 sei die Unterklasse der durch (3.43) definierten
 k-Klasse des gesamten Systems, für die gilt:
 $\text{plim} \sqrt{T}(k_i - 1) = 0$ ($i=1, \dots, N$) und deren Schätzfunktionen
 damit die asymptotische Kovarianzmatrix Q_2 aus (3.44)
 besitzen.

Dann ist K_1 relativ effizient bezüglich K_2 in dem Sinne,
 daß $Q_2 - Q_1$ eine nichtnegativ-definite Matrix ist.

Beweis: Der Beweis kann wie bei Satz 3.7 auf den entspre-
 chenden Beweis für 3SLS zurückgeführt werden. Dieser ist
 bei Schönfeld¹⁾ zu finden.

4. Die k-Klasse als Spezialfall der K-Matrix-Klasse

Nach den vorangegangenen Abschnitten ist klar gewor-
 den, worin der wesentliche Unterschied zwischen k-Klasse
 und K-Matrix-Klasse besteht: Man betrachte nochmals die
 Schätzgleichungen der k-Klasse für das ganze Modell (3.2):

$$(4.1) \quad (Z_1' (I - k_1 \bar{M}) Z_1)_{i,i}^{-1} \hat{\alpha} = (Z_1' (I - k_1 \bar{M}) y_1)_{i,i}^{-1}$$

und vergleiche sie mit den Schätzgleichungen der K-Matrix-
 Klasse:

1) vgl. Schönfeld, P.: Methoden der Ökonometrie, Bd.II,
 a.a.O. S.255.

$$(4.2) \quad \left(\sigma^{ih} z_i' (I - k_{ih} \bar{M}) z_h \right)_{i,h} \hat{\alpha} = \left(\sum_h \sigma^{ih} z_i' (I - k_{ih} \bar{M}) y_h \right)_{i,1}$$

(4.1) erhält man aus (4.2), wenn man $\sigma^{ih}=0$ setzt für $i \neq h$, wenn man also die contemporären Kovarianzen der Störglieder in den Schätzgleichungen nicht beachtet und so tut, als sei Σ diagonal¹⁾. (σ^{ii} kürzt sich links und rechts in (4.1) weg, k_i ($i=1, \dots, N$) ist die Hauptdiagonale der Matrix K in (4.2). Die k -Klasse kann in diesem Sinne als Spezialfall (oder Unterklasse) der K -Matrix-Klasse betrachtet werden.

Auf dieselbe Art ergeben sich die einzelnen LI-Schätzverfahren als Spezialfälle der entsprechenden FI-Schätzverfahren:

Setzt man in der FI/OLS-Schätzfunktion (3.3) bzw. (3.3') $\sigma^{ih}=0$ ($i \neq h$), so erhält man

$$\hat{\alpha} = (z_i' z_i)^{-1}_{i,i} (z_i' y_i)_{i,1} \quad \text{oder}$$

$$\hat{\alpha}_i = (z_i' z_i)^{-1} z_i' y_i \quad (i=1, \dots, N),$$

die OLS-Schätzfunktionen für alle N Strukturgleichungen.

Das gleiche Vorgehen bei der 3SLS-Schätzfunktion (3.7) resultiert in der Schätzfunktion

$$\hat{\alpha} = (z_i' X(X'X)^{-1} X' z_i)^{-1}_{i,i} (z_i' X(X'X)^{-1} X' y_i)_{i,1} \quad \text{oder}$$

$$\hat{\alpha}_i = (z_i' X(X'X)^{-1} X' z_i)^{-1} z_i' X(X'X)^{-1} X' y_i \quad (i=1, \dots, N),$$

also der 2SLS-Schätzfunktion für alle N Strukturgleichungen.

Aus der FIML-Schätzfunktion (3.23) erhält man durch Setzen von $\sigma^{ih}=0$ ($i \neq h$) die LIML-Schätzfunktion (3.24) aus

1) Das bedeutet nicht, daß man Σ a priori als diagonal annimmt.

Definition 2.3' für alle Strukturgleichungen.

Auch aus den DKMC-Schätzgleichungen (3.41) ergeben sich offensichtlich auf die gleiche Weise die Doppel-k-Klasse-Schätzgleichungen für das gesamte Modell. Man betrachte jetzt die Schätzgleichungen (3.40) der h-Vektor-Klasse. Setzt man dort $\sigma^{ik}=0$ ($i \neq k$), so erhält man

$$\left(Z_i' (I - (1 - h_i) h_i) \bar{M} \right) Z_i \hat{\alpha} = \left(Z_i' (I - (1 - h_i) \bar{M}) y_i \right)_{i,1} \quad \text{oder}$$

$$Z_i' (I - (1 - h_i^2) \bar{M}) Z_i \hat{\alpha} = Z_i' (I - (1 - h_i) \bar{M}) y_i \quad (i=1, \dots, N),$$

die Schätzgleichungen der h-Klasse für alle N Gleichungen.

Man bekommt also in der Tat alle LI-Schätzverfahren aus den entsprechenden FI-Schätzverfahren durch nichtbeachten der Kovarianzen der kontemporären Störglieder. Dieses Nichtbeachten hat eine 'größere' asymptotische Kovarianzmatrix zur Folge (vgl. Satz 3.8). Nur wenn Σ wirklich diagonal ist, hat man bei den LI-Schätzverfahren in der Schätzung keine Information vernachlässigt, und alle FI-Schätzverfahren fallen mit ihrem LI-Pendant zusammen.

5. Beziehungen zwischen verschiedenen Schätzverfahren

Im ersten Teil dieses Abschnitts wird eine mathematische Beziehung zwischen 2SLS und der Doppel-k-Klasse¹⁾ auf eine entsprechende zwischen 3SLS und (einer Unterklasse von) DKMC verallgemeinert. Eine ähnliche Beziehung wurde auch von Srivastava²⁾ zwischen 3SLS und GDkC (vgl. Fuß-

1) vgl. Dhrymes, P.J.: *Econometrics*, a.a.O. S.359ff.

2) Srivastava, V.K.: *Three Stage Least-Squares and Generalized Double k-Class Estimators: A Mathematical Relationship*, *International Economic Review* 12, 1971, 312-316.

note 2 auf S.100) hergeleitet. Im zweiten Unterabschnitt wird gezeigt, wie im genau identifizierten Fall alle Schätzfunktionen mit ILS zusammenfallen. Weiter läßt sich (unter Verwendung der Pseudoinversen) zeigen, daß im Falle von zu wenig Beobachtungen (wenn nämlich die Beobachtungszahl T kleiner oder gleich der Anzahl der vorherbestimmten Variablen des gesamten Modells ist) sich alle Schätzverfahren auf die einfachen OLS-Schätzungen reduzieren.

5.1. Eine Beziehung zwischen 3SLS und DKMC

Dhrymes¹⁾ zeigt einen interessanten Zusammenhang zwischen 2SLS und Doppel-k-Klasse-Schätzfunktionen auf. Aus diesem Zusammenhang resultiert, daß jede Schätzfunktion der Doppel-k-Klasse (DkC) aus der 2SLS-Schätzfunktion erzeugt werden kann. Sowohl 2SLS als auch die DkC-Schätzfunktionen können nämlich als OLS-Schätzfunktionen einer jeweils geeignet transformierten Gleichung aufgefaßt werden. Die Gleichung, von der die DkC-Schätzfunktionen abgeleitet werden, kann dann als falsch spezifizierte Version der Gleichung betrachtet werden, von der die 2SLS-Schätzfunktion abgeleitet wird. Ein einfacher Zusammenhang von OLS-Schätzfunktionen, die aus einer falsch und einer richtig spezifizierten Gleichung gewonnen werden, läßt sich jedoch leicht herstellen.

Im folgenden soll nun gezeigt werden, wie man einen ähnlichen Zusammenhang zwischen 3SLS-Schätzfunktion und einigen DKMC-Schätzfunktionen gewinnen kann. Dazu zunächst folgendes Lemma, das eine einfache Verallgemeinerung von

1) vgl. Dhrymes, P.J.: *Econometrics*, a.a.O. S.359ff.

vgl. auch Dhrymes, P.J.: *An Identity between Double k-Class and Two Stage Least Squares*, *International Economic Review* 10, 1969, 114-116.

Lemma 1 (Dhrymes, a.a.O. S.359) darstellt:

Lemma 5.1: Man betrachte das lineare Modell

$$(5.1) \quad y = X\beta + u \quad (y: Tx1, X: TxK, \beta: Kx1, u: Tx1)$$

unter den verallgemeinerten Annahmen

$$E\{u\} = 0, \quad E\{uu'\} = \sigma^2\theta, \quad \theta > 0.$$

X sei aufgeteilt in (X_1, X_2) , β entsprechend in $\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$.

Die GLS-Schätzfunktion von (5.1) ist

$$\hat{\beta} = (X'\theta^{-1}X)^{-1}X'\theta^{-1}y$$

Außerdem sei

$$\tilde{\beta}_1 = (X_1'\theta^{-1}X_1)^{-1}X_1'\theta^{-1}y$$

die Schätzfunktion einer "falsch" spezifizierten Gleichung. (Falsch insofern, als die Variablen aus X_2 außer acht gelassen werden.)

Dann ist

$$(5.2) \quad \tilde{\beta}_1 = \hat{\beta}_1 + S_{12}\hat{\beta}_2,$$

wo $\hat{\beta}_1$, $\hat{\beta}_2$ die entsprechenden Untervektoren von $\hat{\beta}$ bezeichnen

$$\text{und } S_{12} = (X_1'\theta^{-1}X_1)^{-1}X_1'\theta^{-1}X_2.$$

Beweis: Unter Benutzung der partitionierten Matrizen lassen sich die GLS-Schätzgleichungen für $\hat{\beta}$ schreiben als

$$\begin{pmatrix} X_1' \\ X_2' \end{pmatrix} \theta^{-1} (X_1, X_2) \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} X_1' \\ X_2' \end{pmatrix} \theta^{-1} y \quad \text{oder}$$

$$X_1'\theta^{-1}X_1\hat{\beta}_1 + X_1'\theta^{-1}X_2\hat{\beta}_2 = X_1'\theta^{-1}y$$

$$X_2'\theta^{-1}X_1\hat{\beta}_1 + X_2'\theta^{-1}X_2\hat{\beta}_2 = X_2'\theta^{-1}y$$

Löst man den zweiten Teil dieses Gleichungssystems nach $\hat{\beta}_2$ auf, so erhält man

$$\hat{\beta}_2 = (X_2'\theta^{-1}X_2)^{-1}X_2'\theta^{-1}y - (X_2'\theta^{-1}X_2)^{-1}X_2'\theta^{-1}X_1\hat{\beta}_1$$

Weiter ist $X_1'\theta^{-1}y = X_1'\theta^{-1}X_1\tilde{\beta}_1$ nach Definition von $\tilde{\beta}_1$.
Substituiert man nun die soeben erhaltenen Ausdrücke für $\hat{\beta}_2$ und $X_1'\theta^{-1}y$ im ersten Teil des obigen Gleichungssystems, so erhält man

$$X_1'\theta^{-1}X_1\tilde{\beta}_1 = X_1'\theta^{-1}X_1\hat{\beta}_1 - X_1'\theta^{-1}X_2(X_2'\theta^{-1}X_2)^{-1}X_2'\theta^{-1}X_1\hat{\beta}_1 \\ + X_1'\theta^{-1}X_2(X_2'\theta^{-1}X_2)^{-1}X_2'\theta^{-1}y$$

Prämultiplikation mit $(X_1'\theta^{-1}X_1)^{-1}$ ergibt

$$\tilde{\beta}_1 = \hat{\beta}_1 + S_{12}[-(X_2'\theta^{-1}X_2)^{-1}X_2'\theta^{-1}X_1\hat{\beta}_1 + (X_2'\theta^{-1}X_2)^{-1}X_2'\theta^{-1}y] \\ = \hat{\beta}_1 + S_{12}\hat{\beta}_2 \quad \text{q.e.d.}$$

Bemerkung zu Lemma 5.1: In Hinblick auf die späteren Ergebnisse ist es wichtig festzustellen, daß der in Lemma 5.1 dargestellte Zusammenhang zwischen $\tilde{\beta}_1$ und $\hat{\beta}$ eine mathematische Identität darstellt, unabhängig von den stochastischen Voraussetzungen des Modells. So könnte man für θ eine beliebige positiv definite Matrix wählen, die nicht mehr von den stochastischen Annahmen über die Störglieder abhängt, ohne die Aussage von Lemma 5.1 zu verändern. GLS-Schätzung bedeutet hier wie im gesamten Abschnitt 5.1 also immer nur die formale Anwendung der Aitken-Schätzformel (1.3). Insbesondere kann man deshalb natürlich auch keine Aussagen über die stochastischen Eigenschaften der resultierenden Schätzfunktionen machen.

Man betrachte nun wieder das Modell (3.1) bzw. (3.2), zur Abkürzung schreibe man:

$$\hat{Y}_1 = X\hat{\Pi}_1 = X(X'X)^{-1}X'Y_1$$

$$\hat{W}_1 = h_{11}\hat{V}_1 = h_{11}(I - X(X'X)^{-1}X')Y_1 = h_{11}\bar{M}Y_1$$

$$\hat{w}_1 = h_{12}\hat{\phi}_1 = h_{12}\bar{M}y_1$$

mit beliebigen Skalaren h_{11} , h_{12} .

Mit diesen Definitionen erhält man folgendes Lemma:

Lemma 5.2: Die (nichtausführbare¹⁾) 3SLS-Schätzfunktion von α aus (3.2) ist die GLS-Schätzfunktion der Koeffizienten von $\hat{Y} + \hat{W}$ und von x in dem Ansatz

$$(5.3) \quad y + \hat{w} = (\hat{Y} + \hat{W})\gamma + x\beta + \hat{W}\delta + w^*$$

unter der Annahme $E\{w^*w^{*'}\} = \Sigma \otimes I$,

wobei $y = (y_i)_i$, $\hat{w} = (\hat{w}_i)_i$, $\hat{Y} = (\hat{Y}_i)_i$, $\hat{W} = (\hat{W}_i)_i$

$$x = (x_1', \dots, x_N')', \quad \beta = (\beta_1', \dots, \beta_N')', \quad w^* = (w_1^*)_i$$

Beweis: Die GLS-Schätzfunktion der Parameter in (5.3) ist durch folgende Schätzgleichungen definiert:

$$(5.4) \quad \begin{pmatrix} \hat{Y}' + \hat{W}' \\ x' \\ \hat{W}' \end{pmatrix} \theta^{-1} (\hat{Y} + \hat{W}, x, \hat{W}) \begin{pmatrix} \hat{\gamma} \\ \beta \\ \delta \end{pmatrix} = \begin{pmatrix} \hat{Y}' + \hat{W}' \\ x' \\ \hat{W}' \end{pmatrix} \theta^{-1} (y + \hat{w})$$

Multipliziert man dieses Gleichungssystem aus, so erhält man für die einzelnen Untermatrizen:

$$\begin{aligned} (\hat{Y}' + \hat{W}') \theta^{-1} (\hat{Y} + \hat{W}) &= (\hat{Y}' + \hat{W}') (\Sigma^{-1} \otimes I) (\hat{Y} + \hat{W}) = \\ &= \langle \sigma^{ik} (\hat{Y}_i' + \hat{W}_i') (\hat{Y}_k + \hat{W}_k) \rangle_{i,k} = \langle \sigma^{ik} (\hat{Y}_i' \hat{Y}_k + \hat{W}_i' \hat{W}_k) \rangle_{i,k} \end{aligned}$$

$$(\hat{Y}' + \hat{W}') \theta^{-1} x = \langle \sigma^{ik} (\hat{Y}_i' + \hat{W}_i') x_k \rangle_{i,k} = \langle \sigma^{ik} Y_i' x_k \rangle_{i,k}$$

$$(\hat{Y}' + \hat{W}') \theta^{-1} \hat{W} = \langle \sigma^{ik} (\hat{Y}_i' + \hat{W}_i') \hat{W}_k \rangle_{i,k} = \langle \sigma^{ik} \hat{W}_i' \hat{W}_k \rangle_{i,k}$$

$$x' \theta^{-1} x = \langle \sigma^{ik} x_i' x_k \rangle_{i,k}$$

$$x' \theta^{-1} \hat{W} = \langle \sigma^{ik} x_i' \hat{W}_k \rangle_{i,k} = 0, \quad \text{da } x_i' \hat{W}_k = 0$$

$$\hat{W}' \theta^{-1} \hat{W} = \langle \sigma^{ik} \hat{W}_i' \hat{W}_k \rangle_{i,k}$$

Die übrigen Untermatrizen der linken Seite von (5.4) ergeben sich aus einfachen Symmetrieüberlegungen aus den

1) Die Aussagen gelten entsprechend auch für ausführbare Schätzfunktionen.

bisher berechneten. Ausmultiplizieren der rechten Seite von (5.4) ergibt folgende Untervektoren:

$$\begin{aligned} (\hat{Y}' + \hat{W}') \theta^{-1}(\hat{Y} + \hat{W}) &= (\sigma^{ik}(\hat{Y}'_i + \hat{W}'_i))_{i,k} (y_k + \hat{w}_k)_{k,1} = \\ &= (\sum_k \sigma^{ik}(\hat{Y}'_i + \hat{W}'_i)(y_k + \hat{w}_k))_{i,1} \end{aligned}$$

$$X' \theta^{-1}(\hat{Y} + \hat{W}) = (\sum_k \sigma^{ik} X'_i y_k)_{i,1}, \text{ da } X'_i \hat{w}_k = 0$$

$$\hat{W}' \theta^{-1}(\hat{Y} + \hat{W}) = (\sum_k \sigma^{ik} \hat{W}'_i (y_k + \hat{w}_k))_{i,1}.$$

Ausführlich läßt sich (5.4) also schreiben als:

$$\begin{aligned} (\sigma^{ik}(\hat{Y}'_i \hat{Y}'_k + \hat{W}'_i \hat{W}'_k))_{i,k} \hat{\delta} + (\sigma^{ik} Y'_i X_k)_{i,k} \hat{\beta} + (\sigma^{ik} \hat{W}'_i \hat{W}'_k)_{i,k} \hat{\delta} &= \\ (5.5) \quad &= (\sum_k \sigma^{ik}(\hat{Y}'_i + \hat{W}'_i)(y_k + \hat{w}_k))_{i,1} \end{aligned}$$

$$(\sigma^{ik} X'_i Y_k)_{i,k} \hat{\delta} + (\sigma^{ik} X'_i X_k)_{i,k} \hat{\beta} = (\sum_k \sigma^{ik} X'_i y_k)_{i,1}$$

$$(\sigma^{ik} \hat{W}'_i \hat{W}'_k)_{i,k} \hat{\delta} + (\sigma^{ik} \hat{W}'_i \hat{W}'_k)_{i,k} \hat{\delta} = (\sum_k \sigma^{ik} \hat{W}'_i (y_k + \hat{w}_k))_{i,1}.$$

Zieht man nun in (5.5) die dritte von der ersten Gleichung ab, so erhält man folgendes zu (5.5) äquivalente Gleichungssystem:

(5.6)

$$(\sigma^{ik} \hat{Y}'_i \hat{Y}'_k)_{i,k} \hat{\delta} + (\sigma^{ik} Y'_i X_k)_{i,k} \hat{\beta} = (\sum_k \sigma^{ik} \hat{Y}'_i y_k)_{i,1}$$

$$(\sigma^{ik} X'_i Y_k)_{i,k} \hat{\delta} + (\sigma^{ik} X'_i X_k)_{i,k} \hat{\beta} = (\sum_k \sigma^{ik} X'_i y_k)_{i,1}$$

$$(\sigma^{ik} \hat{W}'_i \hat{W}'_k)_{i,k} \hat{\delta} + (\sigma^{ik} \hat{W}'_i \hat{W}'_k)_{i,k} \hat{\delta} = (\sum_k \sigma^{ik} \hat{W}'_i (y_k + \hat{w}_k))_{i,1}$$

$\hat{\gamma}$ und $\hat{\beta}$ können schon aus den ersten beiden Gleichungen von (5.6) bestimmt werden, und man sieht sofort, daß diese ersten beiden (Vektor-) Gleichungen mit den 3SLS-Schätzgleichungen identisch sind, wobei man nur noch beachten muß, daß gewöhnlich die Untervektoren von $\hat{\gamma}$ und $\hat{\beta}$ in anderer Weise zusammengefaßt werden.

q.e.d.

In Analogie zu den bekannten Ergebnissen für Doppel-K-Klasse-Schätzfunktionen läßt sich folgendes Lemma für einige Schätzfunktionen der Doppel-K-Matrix-Klasse formulieren:

Lemma 5.3: Man betrachte die Unterklasse von DKMC, für deren Parametermatrizen K_L und K_R (vgl. Definition 3.5 b)) gilt:

$$K_L = (1 - h_{11} h_{k1})_{i,k} \quad \text{und} \quad K_R = (1 - h_{11} (1 + h_{k2}))_{i,k}$$

mit beliebigen $N \times 1$ -Vektoren $h_1 = (h_{11}, \dots, h_{N1})'$ und $h_2 = (h_{12}, \dots, h_{N2})'$.

Die Schätzfunktionen dieser Unterklasse von DKMC für α aus (3.2) erhält man dann als GLS-Schätzfunktionen¹⁾ der Parameter der Gleichung

$$(5.7) \quad y + \hat{w} = (\hat{Y} + \hat{W})\delta + \lambda\beta + w^*.$$

Bemerkung zu Lemma 5.3: Die in Lemma 5.3 definierte Unterklasse von DKMC könnte man als Doppel-h-Vektor-Klasse (DhVC) bezeichnen. Die h-Vektor-Klasse ergibt sich nämlich (vgl. (3.37)) aus GLS-Schätzung von (5.7), wenn man den Vektor h_2 , der in $\hat{w}$ auf der linken Seite von (5.7) enthalten ist, gleich Null setzt²⁾. Während man also bei

1) Man beachte die Bemerkung zu Lemma 5.1.

2) Daß in (5.7) im Gegensatz zu (3.37) $\hat{Y}$ statt Y steht, darf wegen $\hat{Y}_i' \hat{Y}_k = Y_i' Y_k - \hat{V}_i' \hat{V}_k$ nicht stören. Man erspart sich damit nur die Substitution, die die hVC-Schätzgleichungen in "symmetrische" Form gebracht hat.

der h-Vektor-Klasse (bzw. der h-Klasse) nur die erklärenden abhängigen Variablen um einen proportionalen Anteil der geschätzten reduzierten-Form-Störglieder bereinigt, bereinigt man bei DhVC zusätzlich auch die zu erklärenden abhängigen Variablen. Setzt man in den Schätzgleichungen von DhVC $\sigma^{ik}=0$ ($i \neq k$), so erhält man die Schätzgleichungen der Doppel-k-Klasse des gesamten Modells mit $k_{i1}=(1-h_{i1}^2)$ und $k_{i2}=1-h_{i1}(1+h_{i2})$. Bei beliebig vorgegebenen k_{i1} und k_{i2} läßt sich aus diesen Beziehungen (wenn man komplexe Zahlen zuläßt) immer h_{i1} und h_{i2} berechnen. Das heißt aber, daß das LI-Pendant zur Doppel-h-Vektor-Klasse die Doppel-k-Klasse ist. Neben DKMC ist damit DhVC eine weitere FI-Verallgemeinerung der Doppel-k-Klasse.

Beweis von Lemma 5.3: Nimmt man wieder $E\{w^*w^*\} = \Sigma \otimes I$ an, so lauten die GLS-Schätzgleichungen des Ansatzes (5.7)

$$(5.8) \quad \begin{pmatrix} \hat{y}' + \hat{w}' \\ x' \end{pmatrix} (\Sigma^{-1} \otimes I) (\hat{y} + \hat{w}, x) \begin{pmatrix} \hat{\beta} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} \hat{y}' + \hat{w}' \\ x' \end{pmatrix} (\Sigma^{-1} \otimes I) (y + w)$$

Wie beim Beweis von Lemma 5.2 läßt sich (5.8) ausführlich schreiben als

$$(5.9) \quad \begin{aligned} & (\sigma^{ik} (\hat{y}'_i \hat{y}_k + \hat{w}'_i \hat{w}_k))_{i,k} \hat{\beta} + (\sigma^{ik} \hat{y}'_i x_k)_{i,k} \hat{\beta} = \left(\sum_k \sigma^{ik} (\hat{y}'_i + \hat{w}'_i) (y_k + w_k) \right)_{i,1} \\ & (\sigma^{ik} x'_i y_k)_{i,k} \hat{\beta} + (\sigma^{ik} x'_i x_k)_{i,k} \hat{\beta} = \left(\sum_k \sigma^{ik} x'_i y_k \right)_{i,1} \end{aligned}$$

Es ist aber

$$\hat{y}'_i \hat{y}_k + \hat{w}'_i \hat{w}_k = y'_i y_k - (1-h_{i1}h_{k1}) \hat{y}'_i \hat{y}_k,$$

da $\hat{y}'_i \hat{y}_k = y'_i y_k - \hat{y}'_i \hat{y}_k$ und

$$(\hat{y}'_i + \hat{w}'_i) (y_k + w_k) = (y'_i - [1-h_{i1}(1+h_{k2})] \hat{y}'_i) y_k$$

Setzt man

$$k_{l,ik} = 1 - h_{i1} h_{k1} \quad \text{und}$$

$$k_{r,ik} = 1 - h_{i1} (1 + h_{k2}),$$

so sieht man, daß die Gleichungen (5.9) die Schätzgleichungen der Doppel-h-Vektor-Klasse (DhVC) genannten Unterklasse der Doppel-K-Matrix-Klasse darstellen mit obigen $K_l = (k_{l,ih})_{i,h}$ und $K_r = (k_{r,ih})_{i,h}$, wobei nur noch zu beachten ist, daß gewöhnlich die Untervektoren von $\hat{\gamma}$ und $\hat{\beta}$ in anderer Weise zusammengefaßt werden. q.e.d.

Aus Lemma 5.2 und Lemma 5.3 wird nun klar, daß sich die 3SLS- und DhVC-Schätzfunktionen dadurch unterscheiden, daß das Gleichungssystem (5.7), aus dem die DhVC-Schätzfunktionen abgeleitet werden, eine "falsch" spezifizierte Version des Gleichungssystems (5.3) ist, aus dem die 3SLS-Schätzfunktion gewonnen wird. Einen Zusammenhang zwischen zwei solchen Schätzfunktionen stellt jedoch Lemma 5.1 her. Demnach läßt sich folgender Satz formulieren:

Satz 5.1: Man betrachte das interdependente Simultansystem (3.1) bzw. (3.2). Die 3SLS- und die in Lemma 5.3 und Bemerkung zu Lemma 5.3 definierte DhVC-Schätzfunktionen der Parameter $\begin{pmatrix} \gamma_1 \\ \beta_1 \end{pmatrix}_{1,1}$ sind durch folgende Identität verbunden:

$$(5.10) \quad \begin{pmatrix} \hat{\gamma}(h_1, h_2) \\ \hat{\beta}(h_1, h_2) \end{pmatrix} = \begin{pmatrix} \hat{\gamma} \\ \hat{\beta} \end{pmatrix}_{3SLS} + S \hat{\delta}.$$

Dabei ist $\hat{\gamma} = (\hat{\gamma}_1', \dots, \hat{\gamma}_N')'$, $\hat{\beta} = (\hat{\beta}_1', \dots, \hat{\beta}_N')'$, $\hat{\delta} = (\hat{\delta}_1', \dots, \hat{\delta}_N')'$, $h_1 = (h_{11}, \dots, h_{N1})'$ und $h_2 = (h_{12}, \dots, h_{N2})'$ sind beliebige $N \times 1$ -Vektoren. Die Matrix S in (5.10) ist erklärt durch

$$S = \begin{pmatrix} (\sigma^{ik} (Y_i' Y_k - (1 - h_{i1} h_{k1}) \hat{v}_i' \hat{v}_k))_{i,k} & (\sigma^{ik} Y_i' X_k)_{i,k} \\ (\sigma^{ik} X_i' Y_k)_{i,k} & (\sigma^{ik} X_i' X_k)_{i,k} \end{pmatrix}^{-1} \begin{pmatrix} (\sigma^{ik} h_{i1} h_{k1} \hat{v}_i' \hat{v}_k)_{i,k} \\ 0 \end{pmatrix}$$

und

$$\hat{\delta} = (\sigma^{ik} h_{11} h_{k2} \hat{v}_i \hat{v}_k)^{-1} (h_{11} \sum_k \sigma^{ik} (1+h_{k2}) \hat{v}_i \hat{v}_k)_{i,1} - \hat{\delta}_{3SLS}$$

Beweis: Lemma 5.1, 5.2 und 5.3. Das X_1 von Lemma 5.1 ist hier $X_1 = (\hat{W} + \hat{W}, *)$, das $X_2 = \hat{W}$. Die benötigten Untermatrizen wurden bereits im Beweis von Lemma 5.2 berechnet. $\hat{\delta}$ bekommt man aus dem letzten Teil des Gleichungssystems (5.5).

q.e.d.

Ausgeschrieben und in Kurzform kann man das Gleichungssystem (5.10) auch folgendermaßen darstellen:

$$\hat{\alpha}(h_{11}, h_{12}) = \hat{\alpha}_{3SLS} + (\sigma^{ik} z_i' (I - (1 - h_{11} h_{k1}) \bar{M}) z_k)^{-1}_{i,k} \times \\ \times \left[\left(\sum_k \sigma^{ik} h_{11} (1 + h_{k2}) z_i' \bar{M} y_k \right)_{i,1} - (\sigma^{ik} h_{11} h_{k1} z_i' \bar{M} z_k)_{i,k} \hat{\alpha}_{3SLS} \right]$$

mit α aus (3.2).

5.2. Zusammenhänge zwischen den verschiedenen Schätzverfahren der K-Matrix-Klasse in Extremfällen

Hier sollen die Zusammenhänge zwischen den Schätzverfahren der K-Matrix-Klasse in einigen Extremfällen behandelt werden. Daß alle FI-Schätzverfahren bei diagonalen Kovarianzmatrix Σ der kontemporären Störglieder mit ihrem LI-Pendant zusammenfallen, wurde bereits in Abschnitt 4 gezeigt. Wie bereits angedeutet sind im genau identifizierten Fall 2SLS, 3SLS, LIML und FIML mit ILS identisch. Gezeigt wurde bisher nur (vgl. den Beweis von Satz 2.5), daß in diesem Fall 2SLS und ILS identisch sind. Allgemein kann man folgenden Satz formulieren:

Satz 5.2: Alle N Strukturgleichungen des Systems (3.1) bzw. (3.2) seien genau identifizierbar (d.h. $\bar{R}_i = N_i - 1 = n_i$, $i=1, \dots, N$).

Dann sind die ILS-, 2SLS-, 3SLS-, LIML- und FIML-Schätzfunktionen für α aus (3.2) identisch.

Beweis: 1. Analog zum Beweis für 2SLS (vgl. Beweis von Satz 2.5) wird zunächst gezeigt, daß auch 3SLS mit ILS zusammenfällt.

Im Beweis von Satz 2.5 erhielt man folgenden identischen Ausdruck für die ILS-Schätzgleichungen der i-ten Gleichung (unter der Normierung $\tilde{\gamma}_{ii} = -1$):

$$-(X'X)^{-1}X'y_i + (X'X)^{-1}X'y_i \hat{\gamma}_i = -\hat{\beta}_i.$$

Faßt man alle N Schätzgleichungen zusammen, so erhält man

$$(5.11) \quad \left(-(X'X)^{-1}X'y_i \right)_{i,1} + \left((X'X)^{-1}X'y_i \right)_{i,i} \langle \hat{\gamma}_i \rangle_{i,1} = \left(-\hat{\beta}_i \right)_{i,1}$$

als ILS-Schätzgleichungen für das gesamte Modell. Prämultiplikation von (5.11) mit $(Z_i'X)_{i,i} (\sigma^{ih}I)_{i,h} = (\sigma^{ih}Z_i'X)_{i,h}$, wo I die $R \times R$ -Einheitsmatrix ist, ergibt:

$$\begin{aligned} \left(-\sum_h \sigma^{ih} Z_i'X (X'X)^{-1}X'y_h \right)_{i,1} + \left(\sigma^{ih} Z_i'X (X'X)^{-1}X'y_h \right)_{i,h} \langle \hat{\gamma}_i \rangle_{i,1} &= \\ &= -\left(\sigma^{ih} Z_i'X_h \right)_{i,h} \langle \hat{\beta}_i \rangle_{i,1} \end{aligned}$$

oder durch Zusammenfassung von $\hat{\beta}_i$ und $\hat{\gamma}_i$:

$$\left(\sigma^{ih} Z_i'X (X'X)^{-1}X'y_h \right)_{i,h} \hat{\alpha} = \left(\sum_h \sigma^{ih} Z_i'X (X'X)^{-1}X'y_h \right)_{i,1}.$$

Dies sind genau die 3SLS-Schätzgleichungen.

Wegen der Bedingung $\bar{R}_i = N_i - 1$ (oder äquivalent dazu $R = N_i + R_i - 1$) ($i=1, \dots, N$), ist die Matrix $(\sigma^{ih}Z_i'X)_{i,h}$ eine quadratische $NR \times NR$ -Matrix, die i.a. auch nichtsingulär ist. Durch Prämultiplikation der 3SLS-Schätzgleichungen mit $(\sigma^{ih}Z_i'X)_{i,h}^{-1}$ kann man daher umgekehrt auch von den 3SLS- zu den ILS-Schätzgleichungen kommen. 3SLS- und ILS-Schätzfunktion sind im genau identifizierten Fall also identisch.

2. Als nächstes wird gezeigt, daß FIML und ILS-Schätzfunktion bei genauer Identifizierbarkeit übereinstimmen.

Man betrachte dazu die Log-Likelihood-Funktion der reduzierten Form. Sie lautet

$$(5.12) \quad L = c - \frac{1}{2} \log |\Omega| - \frac{1}{2} \text{tr}(\Omega^{-1} T^{-1} (Y - X\pi)' (Y - X\pi)).$$

Wie man leicht zeigt, nimmt sie ihr absolutes Maximum für

$$\hat{\pi} = (X'X)^{-1} X'Y, \quad \hat{\Omega} = T^{-1} Y' (I - X(X'X)^{-1} X') Y$$

an. Dies sind genau die OLS-Schätzfunktionen für π und Ω . Bei genauer Identifizierbarkeit ist das Gleichungssystem

$$(5.13) \quad B = -\hat{\pi}\Gamma$$

eindeutig nach B und Γ auflösbar, wobei die Nullrestriktionen entsprechend für B und Γ gegeben sind (vgl. 2.1.). Die aus der Auflösung des Gleichungssystems (5.13) resultierende Schätzfunktion $\hat{\alpha}$ für α aus (3.2) ist einerseits nach Definition 2.1 die ILS-Schätzfunktion für α , andererseits kann man auch zeigen¹⁾, daß dieses $\hat{\alpha}$ die Log-Likelihood-Funktion der Strukturform (3.17) absolut maximiert²⁾, also gleichzeitig die FIML-Schätzfunktion für α aus (3.2) darstellt. Maximieren von (5.12) und (3.17) ist demnach über die Beziehung (5.13) bei genauer Identifizierbarkeit äquivalent. Der Grund dafür ist, daß beide Likelihood-Funktionen durch eine nichtsinguläre Transformation miteinander verknüpft sind.

3. Auch LIML und ILS stimmen im genau identifizierbaren Fall überein: Dazu ist zu zeigen, daß die kleinste Wurzel $\hat{\lambda}$ des Eigenwertproblems (2.17) ihr Minimum $\hat{\lambda}=1$ annimmt (vgl. auch 2.3.2.). Als k-Klasse-Schätzfunktion mit $k=\hat{\lambda}=1$ stimmt LIML dann mit 2SLS überein, für das die Äquivalenz mit ILS bereits gezeigt wurde. Auf einen Beweis der

1) vgl. Schönfeld, P.: Methoden der Ökonometrie, Bd. II, a.a.O. S. 285.

2) wobei $\hat{\Sigma} = \hat{\pi}' \hat{\Omega} \hat{\pi}$.

Tatsache, daß in diesem Fall tatsächlich $=1$ ist, wird hier verzichtet. Man findet ihn z.B. bei Lüdeke¹⁾.

q.e.d.

Ein weiterer Extremfall, der neben diagonalem Σ und genauer Identifizierbarkeit noch auftreten kann, ist der folgende: In allen Schätzgleichungen außer OLS und FI/OLS tritt die Inverse der Matrix $X'X$ auf. Damit die Inverse dieser Matrix existiert, ist notwendig, daß

$$\text{rg}(X) = R \leq T \quad .$$

Denn $\text{rg}(X'X) = \text{rg}(X)$ und die $R \times R$ -Matrix $X'X$ wäre singulär, wenn $\text{rg}(X) = T < R$. (Daß die Beobachtungsmatrix X maximalen Rang hat, kann man o.B.d.A. immer annehmen.) Der Fall $\text{rg}(X) = T < R$ tritt jedoch in der Praxis bei größeren Modellen recht häufig auf, da man dort oft sehr viele vorherbestimmte Variablen im Vergleich zu relativ wenigen Beobachtungswerten hat. Man spricht dann von **u n t e r d i m e n s i o n i e r t e n S t i c h p r o b e n**. Um trotzdem zu Schätzungen zu kommen, hat man einige Verfahren entwickelt (meist für 2SLS), insbesondere das Prinzipalkomponentenverfahren²⁾ und ein modifiziertes 2SLS-Verfahren von Theil³⁾. In jüngerer Zeit hat man aber entdeckt, daß das "natürliche" Verfahren im Falle von unterdimensionierten Stichproben die Verwendung von sogenannten Pseudoinversen beinhaltet, "natürlich" in dem Sinne, daß bei überdimensionierten Stichproben (also $\text{rg}(X)=R \leq T$) wieder das

1) vgl. Lüdeke, D.: Schätzprobleme in der Ökonometrie, a.a.O. S.121.

2) Kloeck, T. und L.B.M. Mennes: Simultaneous Equations Estimation Based on Principal Components of Predetermined Variables, *Econometrica* 28, 1960, 45-61. Zu einer Darstellung des Verfahrens auch für 3SLS siehe Friedrich, D.: a.a.O. S. 59ff.

3) Theil, H.: Eine einfache Modifikation der zweistufigen Methode der kleinsten Quadrate für unterdimensionierte Stichproben, IFO-Studien, Berlin/München 1972, S.219-238.

ursprüngliche Verfahren aus der Verwendung dieser Pseudoinversen resultiert.

Die Pseudoinverse ist eine Verallgemeinerung der Inversen für singuläre und sogar nichtquadratische Matrizen. Sie ist (in ihrer einfachsten Form) wie folgt definiert:

Definition 5.1: Gegeben sei eine beliebige Matrix A . Dann heißt die Matrix A^- , die die folgenden vier Bedingungen erfüllt

1. $AA^-A = A$
2. $A^-AA^- = A^-$
3. $(AA^-)' = AA^-$
4. $(A^-A)' = A^-A$

Pseudoinverse von A .

Man zeigt leicht die Existenz und Eindeutigkeit der Pseudoinversen¹⁾. Insbesondere ist die Pseudoinverse für nichtsinguläre Matrizen wegen der Eindeutigkeit gleich der Inversen, da diese alle vier Bedingungen erfüllt.

Verwendet man nun in den DKMC-Schätzgleichungen (3.41) bzw. den DkC-Schätzgleichungen (2.49) die Pseudoinverse $(X'X)^-$ von $X'X$, so erhält man folgenden Satz:

Satz 5.3: Es gelte $R \geq \text{rg}(X) = T \geq \max_i \{n_i + r_i\}$ ²⁾. Die dann in den DKMC- bzw. DkC-Schätzgleichungen (3.41) bzw. (2.49) auftretende nichtexistierende Inverse von $X'X$ ersetze man durch ihre Pseudoinverse $(X'X)^-$ aus Definition 5.1.

Dann fällt die gesamte Doppel-K-Matrix-Klasse mit FI/OLS und die gesamte Doppel-k-Klasse mit OLS zusammen.

1) vgl. Schönfeld, P.: Methoden der Ökonometrie, Bd.II, a.a.O. Anhang C, zu einer umfassenden Darstellung von Pseudoinversen Rao, C.R. und S.K. Mitra: Generalized Inverse of Matrices and its Applications, New York 1971.

2) Diese zweite Ungleichung muß gelten, damit die in OLS auftretende Matrix $Z_1'Z_1$ nichtsingulär ist.

Beweis: Bei Verwendung von Pseudoinversen tritt in allen Unterblöcken der Matrizen rechts und links in den DKMC-Schätzgleichungen (3.41) die Matrix $X(X'X)^{-1}X'$ auf. Prä- und Postmultiplikation von $X(X'X)^{-1}X'$ mit $I = (XX')^{-1}(XX')$ (die Inverse von XX' existiert ja) ergibt:

$$\begin{aligned} X(X'X)^{-1}X' &= (XX')^{-1}X(X'X)(X'X)^{-1}(X'X)X'(XX')^{-1} \\ &= (XX')^{-1}X \quad (X'X) \quad X'(XX')^{-1} \quad (\text{Bed. 1 in Def. 5.1}) \\ &= (XX')^{-1}(XX')(XX')(XX')^{-1} \\ &= I \end{aligned}$$

$X(X'X)^{-1}X'$ reduziert sich also auf die $T \times T$ -Einheitsmatrix, und in den DKMC-Schätzgleichungen ist

$$I - k_{n,ih} \bar{M} = I - k_{n,ih} (I - X(X'X)^{-1}X') = I - k_{n,ih} (I - I) = I \quad (n=1, r; i, h=1, \dots, N).$$

$I - k_{n,ih} \bar{M}$ reduziert sich also auf I , die gesamten DKMC-Schätzgleichungen damit auf die FI/OLS-Schätzgleichungen (3.3').

Mit dem gleichen Argument sieht man, daß sich die DkC-Schätzgleichungen (2.49) auf die OLS-Schätzgleichungen (1.2) reduzieren¹⁾.

q.e.d.

Die in diesem Abschnitt erhaltenen Ergebnisse sind im nächsten Abschnitt tabellarisch zusammengefaßt.

1) Im Fall $\text{rg}(X) = R = T$ kann man auch direkt zeigen:

$$X(X'X)^{-1}X' = XX^{-1}X'^{-1}X' = I.$$

5.3. Schematische Zusammenfassung der wichtigsten Ergebnisse

Aus Übersichtlichkeitsgründen sind in diesem Abschnitt einige wichtige Ergebnisse von Kapitel II schematisch bzw. tabellarisch zusammengefaßt:

Schema 1: Abhängigkeiten zwischen den verschiedenen Klassen von Schätzfunktionen.

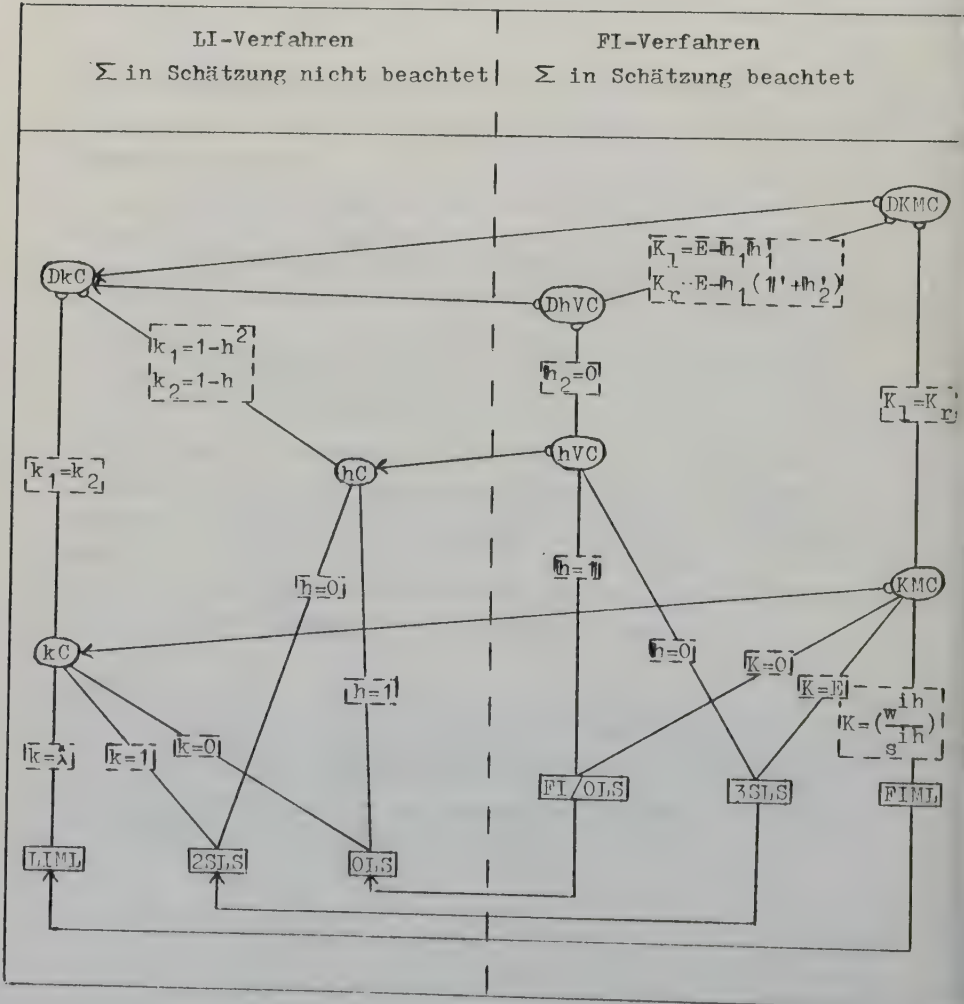


Tabelle 1: Erzeugung der Schätzfunktionen

a) durch Minimierungsprozesse

LI-Verfahren	FI-Verfahren
$OLS = \min \text{tr}(U'U)$	$FI/OLS = \min U'U $
$2SLS = \min \text{tr}(U'X(X'X)^{-1}X'U)$	$3SLS = \min U'X(X'X)^{-1}X'U \quad 1)$
$LIML = \min \frac{\text{tr}(U'U)}{\text{tr}(\hat{V}'\hat{V})}$	$FIML = \min \frac{ U'U }{ \hat{V}'\hat{V} }$

b) als IV-Schätzfunktion

LI-Verfahren	Instr.-Matrix	Instr.-Matrix	FI-Verfahren
OLS	Z_i	$(\Sigma^{-1} \otimes I)Z$	FI/OLS
2SLS	$X(X'X)^{-1}X'Z_i$	$(\Sigma^{-1} \otimes X(X'X)^{-1}X')Z$	3SLS
LIML	$(I - \hat{\lambda}\bar{M})Z_i$	$(\sigma^{ih} (I - \frac{w^{ih}}{s^{ih}} \bar{M}) Z_h)_{i,h}$	FIML

1) Man beachte hier die Fußnote 2) auf Seite 80.

Tabelle 2: Zusammenhänge zwischen den Schätzfunktionen
in einigen Extremfällen

	$rg(X) = R < T$				$R > rg(X) = T > \max_i \{n_i + r_i\}$
	genau identifizierbar		überidentifizierbar		identifizierbar
$\sum$ diagonal	$OLS = FI/OLS$ $ILS = 2SLS = 3SLS = LIML = FIML$		$OLS = FI/OLS$ $2SLS = 3SLS$ $LIML = FIML$		$OLS = DKMC = DkC$
$\sum$ nicht diagonal	OLS $ILS = 2SLS = 3SLS = LIML = FIML$	FI/OLS $LIML = FIML$	OLS $2SLS$ $LIML$	FI/OLS $3SLS$ $FIML$	$OLS = DkC$ $FI/OLS = DKMC$

III. KLEIN-STICHPROBENEIGENSCHAFTEN DER KMC-SCHÄTZFUNKTIONEN

Im Mittelpunkt dieses Kapitels steht die (hier erstmals durchgeführte) Verallgemeinerung eines Verfahrens von Nagar¹⁾ zur approximativen analytischen Behandlung der ersten beiden Klein-Stichproben-Momente der k-Klasse-Schätzfunktionen auf die in Kapitel II behandelten Schätzfunktionen der K-Matrix-Klasse. Die beiden Hauptergebnisse sind in Theorem 4.1 und Theorem 4.2 des Abschnittes 4 zusammengefaßt.

1. Möglichkeiten zur Bestimmung von Klein-Stichprobeneigenschaften von ökonometrischen Schätzfunktionen

Wie Kapitel II zeigte, sind asymptotische Schätzeigenschaften (also Schätzeigenschaften bei unbegrenzt wachsendem Stichprobenumfang T) für die Schätzverfahren der k- und K-Matrix-Klasse relativ einfach zu bestimmen. Dies trifft sicher auf den Nachweis der Konsistenz zu, aber auch über die asymptotische Verteilung der Schätzfunktionen geläufigt man - wenn auch auf mühsamerem Wege - mit Hilfe von zentralen Grenzwertsätzen zu allgemeinen Aussagen (vgl. Satz 2.6 und Satz 3.7 in Kapitel II). Auf Grund dieser asymptotischen Schätzeigenschaften kommt man einerseits zu der Erkenntnis, daß die FI-Schätzverfahren den LI-Schätzverfahren wegen der größeren (asymptotischen) Effizienz vorzuziehen sind, andererseits lassen sich mit Hilfe dieser asymptotischen Kriterien die LI-Schätzverfahren LIML und 2SLS ebenso wenig unterscheiden wie die FI-Schätzverfahren FIML und 3SLS.

Typischerweise kann man jedoch in der Praxis nur auf

1) vgl. Nagar, A.L.: The Bias and Moment Matrix of the General k-Class Estimators..., a.a.O. 575-595.

einen endlichen, meist sogar wegen der wenigen zur Verfügung stehenden Beobachtungen sehr begrenzten Stichprobenumfang T zurückgreifen. Man kann zwar die asymptotischen Ergebnisse als Approximation für die tatsächlichen Schätzeigenschaften betrachten¹⁾, aber natürlich wird die Approximation um so gröber, je kleiner der Stichprobenumfang ist. Wünschenswert wären also analytische Methoden, mit Hilfe derer man zu Aussagen über die Güte einer Schätzfunktion bei endlichem Stichprobenumfang gelangt, mit Hilfe derer man z.B. die Verteilung einer Schätzfunktion bei endlichem T , ihren Erwartungswert oder höhere Momente berechnen könnte. So könnte es durchaus sein, daß 2SLS und LIML-Schätzfunktion auf Grund von "Klein-Stichproben"-Eigenschaften (d.h. von Eigenschaften, die für endliches T berechnet wurden) sich unterscheiden lassen in dem Sinne, daß die eine Schätzfunktion günstigere Schätzeigenschaften als die andere aufweist, dieser also bei der Schätzung vorzuziehen wäre. Es ist auch nicht ausgeschlossen, daß sogar OLS trotz der Inkonsistenz bei kleinem T den k -Klasse-Verfahren mit $\text{plim } k = 1$ vorzuziehen ist.

Die mathematisch-analytische Behandlung von Klein-Stichprobeneigenschaften der ökonometrischen Schätzfunktionen ist jedoch äußerst schwierig, in vielen Fällen sogar unmöglich. Trotz der vielen Veröffentlichungen zu diesem Thema (vgl. den nächsten Abschnitt) wurde bisher noch kein allgemein anwendbares Verfahren gefunden, das es erlauben würde, endliche Stichprobeneigenschaften zu bestimmen. Zwar kann man in einigen Spezialfällen (als einfachsten Fall dieser Art betrachte man den Nachweis der BLUE-Eigenschaft der OLS-Schätzfunktion bei nichtstochastischen

1) vgl. dazu Schönfeld, P.: Methoden der Ökonometrie, Bd.I, Berlin und Frankfurt 1969, Abschnitt 6.5.

erklärenden Variablen (II,1.1.)) zu exakten Ergebnissen gelangen, im allgemeinen (in der Praxis meist allein interessierenden) Fall ist dies meistens nicht möglich, da zentrale Grenzwertsätze oder z.B. die günstigen Eigenschaften des plim-Operators im endlichen Fall nicht mehr anwendbar sind. Oft weicht man dann auf approximative Verfahren aus, die z.B. die Klein-Stichproben-Momente einer Schätzfunktion nicht exakt sondern nur angenähert bestimmen. Auf der anderen Seite sind wirklich exakte Verfahren zur Bestimmung von Klein-Stichprobeneigenschaften oft auf wenige Spezialfälle beschränkt (z.B. Beschränkung auf nur drei Gleichungen oder auf zwei endogene Variablen in einer Gleichung u.ä.). Eine andere Unterscheidungsmöglichkeit von Verfahren zur Bestimmung von Klein-Stichprobeneigenschaften ist neben dem Unterscheidungsmerkmal approximatives-exaktes Verfahren die Unterteilung in Verfahren, die versuchen (meist unter der Voraussetzung der Normalverteilung der Störglieder), die Verteilungsfunktion (oder Dichtefunktion) einer Schätzfunktion zu bestimmen und in Verfahren, die die Momente (Bias, Momentenmatrix) einer Schätzfunktion zu bestimmen versuchen. Im nächsten Abschnitt folgt ein kurzer Überblick über die Literatur (der keinerlei Anspruch auf Vollständigkeit erhebt.)

2. Literaturhinweise

Die ersten exakten Klein-Stichprobenergebnisse für OLS und 2SLS gehen auf Basmann¹⁾, Bergstrom²⁾ und

-
- 1) Basmann, R.L.: A Note on the Exact Finite Sample Frequency Functions of Generalized Classical Linear Estimators in two Leading Over-Identified Cases, JASA 56, 1961, 619-636.

Basmann, R.L.: A Note on the Exact Finite Sample Frequency Functions of Generalized Classical Linear Estimators in a Leading Three Equation Case, JASA 58, 1963, 161-171.

Basmann, R.L.: Remarks Concerning the Application of Exact Finite Sample Distribution Functions of GCL Estimators in Econometric Statistical Inference, JASA 58, 1963, 943-976.

- 2) Bergstrom, A.R.: The Exact Sampling Distribution of Least Squares and Maximum Likelihood Estimators of the Marginal Propensity to Consume, E 30, 1962, 480-490.

Kabe¹⁾ zurück (zur Abkürzung der Zeitschriften vergl. das Literaturverzeichnis). Sie berechneten die exakte Klein-Stichproben-Dichtefunktion der 2SLS-bzw. OLS-Schätzfunktion in speziellen Modellen mit höchstens drei Gleichungen. Weitere Arbeiten in dieser Richtung sind die von Richardson²⁾, Takeuchi³⁾, Sawa⁴⁾ und Ullah/Nagar⁵⁾. Bis auf die Arbeit von Ullah/Nagar, die sich auf drei in einer Gleichung auftretende endogene Variablen beschränkt, gelten diese Ergebnisse nur für den Fall von zwei in einer Gleichung auftretende endogene Variable. Sie enthalten auch Ergebnisse über die LIML-Schätzfunktion. Außerdem bestätigen sie (innerhalb des behandelten Spezialfalles) die bekannte Basman'sche Vermutung, daß die Momente der 2SLS-Schätzfunktion existieren, sofern ihre Ordnung kleiner als $N_i + \bar{R}_i$ ist (N_i ist die Zahl der in der i-ten Gleichung auftretenden endogenen Variablen, $\bar{R}_i$ die Anzahl der aus der i-ten Gleichung ausgeschlossenen vorherbestimmten Variablen). Für eine beliebige Anzahl von eingeschlossenen

1) Kabe, D.G.: A Note on the Exact Distributions of the GCL Estimators in two Leading Overidentified Cases, JASA 58, 1963, 535-537.

Kabe, D.G.: A Note on the Exact Distributions of the GCL Estimators in a Leading Three-Equation Case, JASA 59, 1964, 881-894.

2) Richardson, D.H.: The Exact Distribution of a Structural Coefficient Estimator, JASA 63, 1968, 1214-1226.

3) Takeuchi, K.: Exact Sampling Moments of the Ordinary Least Squares, Instrumental Variable, and Two-Stage Least Squares Estimators, IER 11, 1970, 1-12.

4) Sawa, T.: The Exact Sampling Distribution of Ordinary Least Squares and Two Stage Least Squares Estimators, JASA 64, 1969, 923-937.

5) Ullah, A. und A.L. Nagar: The Exact Mean of the Two-Stage Least Squares Estimator of the Structural Parameters in an Equation having Three Endogeneous Variables, E 42, 1974, 749-758.

endogenen Variablen wird dieses Ergebnis von Mariano¹⁾ verallgemeinert, sofern nur nichtstochastische vorherbestimmte Größen auftreten. Exakte Klein-Stichproben-Ergebnisse (wieder in einem anderen Spezialfall) findet man bei Sawa²⁾ für allgemeine k-Klasse-Schätzfunktionen. Insbesondere existieren keine Momente für $k > 1$.

Verschiedene Verfahren zur Approximation der Dichte- oder Verteilungsfunktionen (bzw. deren Momente) von k-Klasse- (oder auch Instrumenten-Variablen-) Schätzfunktionen findet man z.B. bei Sargan/Mikhail³⁾, Kadane⁴⁾ ('small ϵ -asymptotics'), Mariano⁵⁾ ('large μ -asymptotics'), Brown/Kadane/Ramage⁶⁾ und Anderson/Sawa^{7),8)}.

-
- 1) Mariano, R.S.: The Existence of Moments of the Ordinary Least Squares and Two-Stage Least Squares Estimators, E 40, 1972, 643-652.
 - 2) Sawa, T.: Finite-Sample Properties of the k-Class Estimators, E 40, 1972, 653-679.
 - 3) Sargan, J.D. und W.M. Mikhail: A General Approximation to the Distribution of Instrumental Variables Estimates, E 39, 1971, 131-169.
 - 4) Kadane, J.B.: Comparison of k-Class Estimators when the Disturbances are Small, E 39, 1971, 723-737.
 - 5) Mariano, R.S.: Approximations to the Distribution Functions of the Ordinary Least Squares and Two-Stage Least Squares Estimators in the Case of Two Included Endogeneous Variables, E 41, 1973, 67-77.
Mariano, R.S.: Approximations to the Distribution Functions of Theil's k-Class Estimators, E 41, 1973, 715-721.
Mariano, R.S.: Some Large-Concentration-Parameter Asymptotics for the k-Class Estimators, JE 3, 171-177, 1975.
 - 6) Brown, G.F., Kadane, J.B. und J.G. Ramage: The Asymptotic Bias and Mean-Squared Error of Double k-Class Estimators when the Disturbances are Small, IER 15, 1974, 667-679.
 - 7) Anderson, T.W. und T. Sawa: Distributions of Estimates of Coefficients of a Single Equation in a Simultaneous System and their Asymptotic Expansions, E 41, 1973, 685-716.
 - 8) An weiteren Arbeiten sind zu nennen: Quandt, R.: On Certain Small Sample Properties of k-Class Estimators, IER 6, 1965, 92-104, sowie Mariano, R.S. und T. Sawa: The exact Finite-Sample Distribution of the Limited-Information Maximum Likelihood Estimator in the Case of Two Included Endogeneous Variables, JASA 67, 1972, 159-163.

Die erste Arbeit in Richtung Approximation der Momente der k-Klasse-Schätzfunktionen (speziell 2SLS) ist jedoch der Artikel von Nagar¹⁾. Während man die asymptotische Kovarianzmatrix (vgl. II, Satz 2.6) unter gewissen Voraussetzungen als die Approximation an die Klein-Stichproben-Kovarianzmatrix betrachten kann, die nur den Anteil mit der Wachstumsordnung T^{-1} berücksichtigt, berücksichtigt Nagars Approximation auch Anteile bis zur $O(T^{-2})$ (bzw. beim Bias bis zur $O(T^{-1})$), stellt also eine bessere Approximation dar. Natürlich wächst auch hier die Güte der Approximation mit wachsendem T . Eine genaue Darstellung von Nagars Approximations-Methode folgt im nächsten Abschnitt.

Eine Verallgemeinerung von Nagars oben genannten Resultat auf die Doppel-k-Klasse ist ebenfalls von Nagar ausgeführt²⁾. Eine weitere Verallgemeinerung (in anderer Richtung) stellt der Artikel von Nagar/Gupta³⁾ dar: In ihm wird die Momentenmatrix zur $O(T^{-2})$ der 2SLS-Schätzfunktion von zwei verschiedenen Gleichungen eines interdependenten Modells analysiert. In einer neueren Arbeit gibt Sargan⁴⁾ Bedingungen für die Gültigkeit von Nagars Approximation an.

-
- 1) Nagar, A.L.: The Bias and Moment Matrix of the General k-Class Estimators of the Parameters in Simultaneous Equations, E 27, 1959, 575-595.
 - 2) Nagar, A.L.: Double k-Class Estimators of Parameters in Simultaneous Equations and their Small Sample Properties, IER 3, 1962, 168-188.
 - 3) Nagar, A.L. und Y.P. Gupta: The Moment Matrix of the Two-Stage Least-Squares Estimator of Coefficients in Different Equations of a Complete System of Simultaneous Equations, E 38, 1970, 39-49.
 - 4) Sargan, J.D.: The Validity of Nagar's Expansion for the Moments of Econometric Estimators, E 42, 1974, 169-176.

In der vorliegenden Arbeit werden nun Nagars Ergebnisse auf den Fall der in Kapitel II behandelten K-Matrix-Klasse-Schätzfunktionen (speziell der (nichtausführbaren) 3SLS-Schätzfunktion) verallgemeinert: Bias zur $O(T^{-1})$ und Momentenmatrix zur $O(T^{-2})$ der KMC-Schätzfunktionen werden analysiert. Dieses Ergebnis stellt eines der bislang sehr wenigen Ergebnisse¹⁾ über Klein-Stichprobeneigenschaften von Full-Information-Schätzfunktionen dar.

3. Nagars Verfahren zur Bestimmung des Bias und der Momentenmatrix der k-Klasse-Schätzfunktionen

Nagars Verfahren²⁾ ist im Grunde genommen recht nahelegend, da es sich zunächst am Beweis der Konsistenz orientiert. Der Stichprobenfehler $\hat{\alpha}_i(k) - \alpha_i$ der durch die Schätzgleichungen I, (2.39) bestimmten k-Klasse-Schätzfunktion $\hat{\alpha}_i(k)$ läßt sich (entsprechend dem Beweis von I, Satz 1.2) ausdrücken durch die Schätzfunktion selbst, wo nur y_i durch das Störglied u_i ersetzt ist, also:

$$(3.1) \quad e_i(k) = \hat{\alpha}_i(k) - \alpha_i = (Z_i'(I - k\bar{M})Z_i)^{-1} Z_i'(I - k\bar{M})u_i.$$

Ersetzt man in (3.1) das in Z_i enthaltene Y_i durch den entsprechenden reduzierten-Form-Anteil $X\Pi_i + V_i$ und beachtet den Zusammenhang $V_i = UD_i$, wo D_i ein entsprechender Teil von Γ^{-1} ist, so läßt sich der Stichprobenfehler $e_i(k)$ aus (3.1) darstellen als Funktion der Störglieder und ansonsten nur nichtstochastischen Größen (nämlich dem als

1) vgl. Sargan, J.D.: The Finite Sample Distribution of FIML Estimators, paper presented at the World Congress of the Econometric Society, 1970.

2) Nagar, A.L.: The Bias and Moment Matrix..., a.a.O. 575-595.

nichtstochastisch vorausgesetzten X sowie den wahren Parametern). Bis hierhin ist offensichtlich das Vorgehen wie das beim Beweis der Konsistenz (oder Inkonsistenz). Statt des plim -Operators muß man nun jedoch, um den Bias zu berechnen, den Erwartungswert-Operator E auf die rechte Seite von (3.1) anwenden. Während man den plim ohne weiteres in die Inverse auf der rechten Seite von (3.1) ziehen kann (dies beruht auf dem bekannten Satz von Slutsky¹⁾), ist dies mit dem Erwartungswert-Operator als linearem Operator nicht mehr möglich. Exakte Ergebnisse wird man also mit dieser direkten Methode nicht erwarten dürfen.

Man kann aber die Inverse in (3.1) so umordnen, daß sie in der Form

$$(3.2) \quad (I-S)^{-1}$$

mit einer hier nicht näher spezifizierten Matrix S dargestellt werden kann. Für die Matrix (3.2) existiert eine unendliche Reihenentwicklung der Form²⁾

$$(3.3) \quad (I-S)^{-1} = \sum_{n=0}^{\infty} S^n$$

- Entwicklung (3.3) existiert, wenn die Absolutwerte der charakteristischen Wurzeln von S kleiner als 1 sind. Sie stellt offensichtlich eine Matrix-Verallgemeinerung einer skalaren geometrischen Reihe dar.

Vernachlässigt man in (3.3) alle Summanden, deren Wachstumsordnung³⁾ kleiner als T^{-1} ist (also Glieder der Ordnung T^{-2} , T^{-3} , ...), so erhält man eine Approximation des Stichprobenfehlers $e_i(k)$. Wendet man dann gliedweise den

1) vgl. Wilks, S.S.: Mathematical Statistics, New York 1962, S.102-103.

2) vgl. Gröbner, W.: Matrizenrechnung, Mannheim 1966, S.207.

3) zum Begriff der Wachstumsordnung vgl. Anhang A1.

Erwartungswertoperator E auf diese Approximation für $e_i(k)$ bzw. auf die entsprechende Approximation für $e_i(k)e_i(k)'$ an, so bekommt man Nagars Resultat¹⁾ für den Bias zur $O(T^{-1})$ und die Momentenmatrix (eigentlich: die Matrix des mittleren quadratischen Fehlers (mean-squared error matrix)) zur $O(T^{-2})$ der k -Klasse-Schätzfunktionen, für die noch vorausgesetzt ist, daß $k = 1 + T^{-1}\alpha$, sich k also nur um einen (nichtstochastischen) Term der $O(T^{-1})$ von 1 (und damit der 2SLS-Schätzfunktion) unterscheidet. Dies impliziert insbesondere die Konsistenz. Wie Nagar weiter zeigt, kann man einen Wert für α finden, der eine zur $O(T^{-1})$ erwartungstreue Schätzfunktion erzeugt, sowie einen anderen Wert für α , der die Determinante der Momentenmatrix zur $O(T^{-2})$ minimiert. Hier wird also die Flexibilität in der Wahl des Parameters k ausgenutzt, um zu günstigen (approximativen) Klein-Stichprobeneigenschaften zu gelangen.

Wie man sieht, müssen einige Einschränkungen gemacht werden, damit Nagars Approximation gültig ist: Die Matrix X der Vorherbestimmten muß nichtstochastisch sein, d.h. vor allem, daß man sich auf statische Systeme beschränken muß. Die Entwicklung (3.3) muß gültig sein. Wie bereits erwähnt ist dies der Fall, wenn die Absolutwerte der charakteristischen Wurzeln von S kleiner als 1 sind. Da der Erwartungswert als Integral definiert ist, bedeutet die gliedweise Anwendung des Erwartungswertoperators auf die Entwicklung (3.3) die gliedweise Integration einer unendlichen Reihe, die nur unter gewissen Bedingungen möglich ist. Vor allem muß der Integrand integrierbar sein, was bedeutet, daß die ersten und zweiten Momente existieren müssen. Nagars Approximation beweist also nicht die Existenz von Klein-Stichproben-Bias und -Momentenmatrix,

1) Nagar, A.L.: The Bias and..., a.a.O. Theorem I und Theorem II.

sondern stellt nur eine Approximation dar, falls diese existieren. Bedingungen für die Existenz sind bei Sargan¹⁾ angegeben.

Trotz dieser Einschränkungen ist Nagars Verfahren von Interesse, vor allem, weil weder die Zahl der Gleichungen noch die Zahl der Variablen in irgendeiner Weise beschränkt werden müssen. Nagars Verfahren ist daher gerade auch auf (in der Praxis meist vorkommende) größere Modelle anwendbar.

Während in Nagar und Guptas Artikel²⁾ das oben beschriebene Verfahren auf die Berechnung der Momentenmatrix der 2SLS-Schätzfunktion von zwei verschiedenen Gleichungen eines vollständigen Systems verallgemeinert wird, wird in den folgenden Abschnitten Nagars Verfahren auf die KMC-Schätzfunktionen von II, Definition 3.4 erweitert.

4. Der Bias und die Momentenmatrix der K-Matrix-Klasse-Schätzfunktionen

Aus schreibtechnischen Gründen müssen zunächst einige Matrizen neu definiert werden, teilweise im Gegensatz zu Kapitel I und II: In diesem und den folgenden Abschnitten³⁾ ist

$$Z_i = (X\Pi_i, X_i) , \quad \bar{Z}_i = (Y_i, X_i)$$

$$N^* = X(X'X)^{-1}X'$$

$$(4.1) \quad \bar{V}_i = U\bar{D}_i , \text{ wobei } \bar{D}_i \text{ der zur } i\text{-ten Gleichung gehörende Teil von } -\Gamma^{-1} \text{ ist } (Y_i = X\Pi_i + \bar{V}_i),$$

$$V_i = U D_i \text{ mit } D_i = (\bar{D}_i, 0), 0 \text{ ist } N \times r_i,$$

1) vgl. Sargan, J.D.: The Validity of Nagar's Expansion..., a.a.O. 169-176.

2) Nagar, A.L. und Y.P. Gupta: a.a.O. 39-49.

3) Sofern hier Matrizen auftreten, die nicht neu definiert sind, haben sie dieselbe Bedeutung wie in Kapitel I und II.

$$\Omega_{ih} = D_i' \Sigma D_h$$

$$(4.1) \quad q_{ih} = D_i' \sigma_h \quad (\sigma_h \text{ ist die } h\text{-te Spalte von } \Sigma)$$

$$Q = (q_{ih})_{i,h} = (\sigma^{ih} z_i' z_h)^{-1}_{i,h} \quad (\sigma^{ih} \text{ ist das } (i,h)\text{-te Element von } \Sigma^{-1})$$

Die Schätzfunktionen der K-Matrix-Klasse (vgl. II, (3.32)) lassen sich dann schreiben als

$$(4.2) \quad \hat{\alpha}(K) = (\sigma^{ih} (\bar{z}_i' \bar{z}_h - k_{ih} \hat{v}_i' \hat{v}_h))^{-1}_{i,h} (\sum_h \sigma^{ih} (\bar{z}_i' \bar{z}_h - k_{ih} \hat{v}_i' \hat{v}_h) y_h)_{i,1}$$

mit $\alpha' = (\alpha_1', \dots, \alpha_N')$ aus II, (3.2), der $N \times N$ -Matrix $K = (k_{ih})$

sowie $\hat{v}_1 = (\hat{v}_1, 0) = ((I - X(X'X)^{-1}X')Y_1, 0)$.

Im Folgenden werden nur die nichtausführbaren KMC-Schätzfunktionen betrachtet, daher ist

$$(4.3) \quad s^{ih} = \sigma^{ih} \text{ nichtstochastisch.}$$

In den folgenden Abschnitten wird meist s anstelle des gleichbedeutenden σ geschrieben. Weiter wird analog zu Nagars Entwicklung für die k -Klasse die folgende Annahme über die Matrix K gemacht: Bis auf einen nichtstochastischen Ausdruck der Ordnung T^{-1} soll sie sich nicht von E (der Matrix, die nur aus Einsen besteht) unterscheiden, d.h.

$$(4.4) \quad K = E + T^{-1}\mathcal{X}$$

mit nichtstochastischer $N \times N$ -Matrix $\mathcal{X} = (\mathcal{x}_{ih})$, die symmetrisch sei: $\mathcal{X} = \mathcal{X}'$.

Insbesondere ist durch (4.4) gewährleistet, daß die Schätzfunktionen konsistent sind (vgl. II, Satz 3.5). Nach II, Satz 3.7 haben sie sogar dieselbe asymptotische Verteilung wie 3SLS (wegen $\text{plim} \sqrt{T}(K-E) = \lim \sqrt{T} T^{-1} \mathcal{X} = \lim T^{-1/2} \mathcal{X} = 0$).

Über die Störglieder $u(t)$ wird zusätzlich zu Annahme I,1 vorausgesetzt, daß sie einer N-dimensionalen Normalverteilung folgen:

$$(4.5) \quad u(t) \sim N(0, \Sigma).$$

Außerdem soll für die Matrix der Vorherbestimmten gelten:

$$(4.6) \quad X \text{ ist nichtstochastisch.}$$

Nach dem im vorangehenden Abschnitt beschriebenen Verfahren erhält man unter den oben angegebenen Annahmen und Definitionen den Bias und die Momentermatrix der KMC-Schätzfunktionen. Die Ergebnisse sind in den folgenden beiden Theoremen zusammengefaßt. Der Beweis folgt in Abschnitt 5.

Theorem 4.1: Unter den allgemeinen Annahmen über das lineare interdependente Modell aus Kapitel I, sowie unter den Annahmen (4.3), (4.4), (4.5) und (4.6), ist der Bias der KMC-Schätzfunktionen (4.2) zur $O(T^{-1})$ gleich:

$$\begin{aligned} E\{e(K)\} &= E\{\hat{a}(K)\} - \alpha = \\ &= Q \left[\left(\sum_h s^{ih} (R - (n_h + r_h) - x_{ih}) q_{ih} \right)_{i,1} \right. \\ &\quad - \left(\sum_{\substack{h,k \\ h \neq k}} s^{ih} \text{tr} \left(\sum_j s^{jk} q_{hj} z_j' z_h \right) q_{ik} \right)_{i,1} \\ &\quad \left. - \left(\sum_{h,k} s^{ih} \sum_j s^{jk} z_i' z_j q_{jh} q_{hk} \right)_{i,1} \right] \end{aligned} \quad (4.7)$$

Die Summation geht dabei immer von $1, \dots, N$ und tr bezeichnet die Spur einer Matrix.

Theorem 4.2: Unter den allgemeinen Annahmen über das lineare interdependente Modell aus Kapitel I, sowie unter den Annahmen (4.3), (4.4), (4.5) und (4.6), ist die Momentenmatrix der KMG-Schätzfunktionen (4.2) zur $O(T^{-2})$ gleich:

$$(4.8) \quad E\{e(K)e(K)'\} = E\{(\hat{A}(K) - \alpha)(\hat{A}(K) - \alpha)'\} =$$

$$Q + Q(A(\mathcal{X}) + B(\mathcal{X}) + B(\mathcal{X})' + C + D + C')Q$$

$A(\mathcal{X})$ und $B(\mathcal{X})$ bedeutet, daß diese Matrizen von der Matrix $\mathcal{X}$, die in (4.4) definiert ist, abhängen. Die Matrizen in (4.8) sind wie folgt definiert:

$$Q = (\sigma_{ih} z_i' z_h)^{-1}_{i,h}, \text{ der einzige Term der } O(T^{-1}) \text{ wurde}$$

bereits in (4.1) definiert.

$$A(\mathcal{X}) = \left(\sum_{h,k} s^{ih} s^{jk} (R_{q_{ik} q'_{jh}} + (R - \alpha_{ih})(R - \alpha_{jk}) q_{ih} q'_{jk} + R \sigma_{hk} \Omega_{ij}) \right)_{i,j}$$

$$\begin{aligned} B(\mathcal{X}) = & - \left(\sum_{h,k,j,l} s^{ih} s^{jk} s^{lm} \times \right. \\ & \times [z_i' z_j z_{jl} (q_{lh} q'_{mk} + q_{lk} q'_{mh} + (R - \alpha_{lm}) \sigma_{hk} \Omega_{lm}) + \\ & z_i' z_m (q_{jh} q'_{jl} q_{lk} + (R - \alpha_{jk}) q'_{jk} q_{jl} q_{lh} + \sigma_{hk} \text{tr}(q_{jl} \Omega_{lj})) + \\ & z_i' z_l q_{lj} (q_{jh} q'_{mk} + (R - \alpha_{jk}) q'_{jk} q'_{mh} + \sigma_{hk} \Omega_{jm}) + \\ & (q_{ik} q'_{lh} + (R - \alpha_{ih}) q_{ih} q'_{lk} + \sigma_{hk} \Omega_{il}) q_{lj} z_j' z_m + \\ & \left. \text{tr}(z_j z_{jl} z_l') (q_{ik} q'_{mh} + (R - \alpha_{ih}) q_{ih} q'_{mk} + \sigma_{hk} \Omega_{lm}) \right] \Big)_{i,m} \\ C = & \left(\sum_{h,k,j,l,m,p} s^{ih} s^{jk} s^{lm} s^{pq} z_i' \times \right. \\ & \times [(q_{ph} q'_{pm} z_m' z_j z_{jl} q_{lk} I + z_m q_{mp} q_{pk} q'_{lh} q_{lj} z_j' + \sigma_{hk} z_j z_{jl} \Omega_{lp} q_{pm} z_m') z_q + \\ & (z_m q_{mp} z_p' z_j z_{jl} q_{lh} q'_{lk} + z_p q_{pm} z_m' z_j z_{jl} q_{lk} q'_{qh} + \sigma_{hk} z_j z_{jl} \Omega_{lp} \text{tr}(z_m q_{mp} z_p')) + \\ & (z_l q_{lj} z_j' q_{mh} q'_{mp} q_{pk} + \text{tr}(z_j z_{jl} z_l') q_{mk} q_{mp} q_{ph} I + \sigma_{hk} z_j z_{jl} z_l' \text{tr}(q_{mp} \Omega_{pm})) z_q + \\ & \left. (z_l q_{lj} z_j' z_p q_{pm} q'_{mh} q'_{lk} + \text{tr}(z_j z_{jl} z_l') z_p q_{pm} q_{mk} q'_{qh} + \sigma_{hk} z_j z_{jl} z_l' z_p q_{pm} \Omega_{mq}) \right] \Big)_{i,q} \end{aligned}$$

$$\begin{aligned} \hat{D} = & \left(\sum_{h,k,j,l,m,p} s^{il} s^{jh} s^{mk} s^{pq} \right) \times \\ & \times \left[(q_{ih} q'_{pk} Q_{pm} Z'_m \text{tr}(Q_{lj} Z'_j Z_l) + q_{ik} q'_{ph} Q_{pm} Z'_m Z_l Q_{lj} Z'_j + \epsilon_{hk} \Omega_{lp} Q_{pm} Z'_m Z_j Q_{jl} Z'_l) Z_q \right. \\ & (q_{ih} q'_{qk} \text{tr}(Q_{lj} Z'_j Z_l) \text{tr}(Q_{mp} Z'_p Z_m) + q_{ik} q'_{qh} \text{tr}(Z_l Q_{lj} Z'_j Z_p Q_{pm} Z'_m) + \\ & \left. \epsilon_{hk} \Omega_{lq} \text{tr}(Z_l Q_{lj} Z'_j Z_m Q_{mp} Z'_p)) + \right. \\ & Z'_i (Z_j Q_{jl} q_{lh} q'_{pk} Q_{pm} Z'_m + Z_m Q_{mp} q_{ph} q'_{lk} Q_{lj} Z'_j + \epsilon_{hk} \text{tr}(Z_m Q_{mp} \Omega_{pl} Q_{lj} Z'_j) I) Z_q + \\ & Z'_i (Z_j Q_{jl} q_{lh} q'_{qk} \text{tr}(Z_m Q_{mp} Z'_p) + Z_m Q_{mp} Z'_p Z_j Q_{jl} q_{lk} q'_{qh} + \\ & \left. \left. \epsilon_{hk} Z_p Q_{pm} Z'_m Z_j Q_{jl} \Omega_{lq} \right) \right]_{i,q} \end{aligned}$$

Korollar: Setzt man in (4.7) bzw. (4.8) $\epsilon_{ih} = 0$ ($i, h=1, \dots, N$), so erhält man den Bias zur $O(T^{-1})$ und die Momentenmatrix zur $O(T^{-2})$ der (nichtausführbaren) 3SLS-Schätzfunktion.

Bemerkungen zu Theorem 4.1 und Theorem 4.2:

Theorem 4.1 beinhaltet Nagars Ergebnis über den Bias der k-Klasse-Schätzfunktionen als Spezialfall: Dies sieht man leicht folgendermaßen ein: Die k-Klasse-Schätzfunktionen einer bestimmten (der i-ten) Gleichung kann man auch auffassen als KMC-Schätzfunktion eines Modells mit nur einer Gleichung, d.h. mit $N=1$. Daher bestehen in diesem Fall alle Summen in (4.7) nur aus einem Summanden, alle partitionierten Matrizen und Vektoren reduzieren sich auf eine (die (i,i)-te) Untermatrix und auf einen (den i-ten) Untervektor. Außerdem reduziert sich Σ auf σ_{ii} und Q in (4.7) auf $Q_{ii} = (\sigma^{ii} Z'_i Z_i)^{-1} = \sigma_{ii} (Z'_i Z_i)^{-1}$. Der erste Summand in (4.7) ergibt

$$\begin{aligned} & \sigma^{ii} (R - (n_i + r_i) - x_{ii}) q_{ii} = \\ & \sigma^{ii} (R - (n_i + r_i) - x_{ii}) D'_i \sigma_i \\ & \sigma^{ii} (R - (n_i + r_i) - x_{ii}) d_i \sigma_{ii}, \text{ da } \Sigma \text{ und damit auch die } i\text{-te} \\ & \text{Spalte von } \Sigma \text{ nur aus } \sigma_{ii} \text{ besteht. Der zweite Summand in} \end{aligned}$$

(4.7) wird Null, da nur über $h \neq k$ summiert wird, in unserem Fall aber immer $h=k=i$ ist. Der dritte Summand in (4.7) ergibt

$$\begin{aligned} \sigma_{i1}^{ii} \sigma_{i1}^{ii} Z_i' Z_i \sigma_{i1} (Z_i' Z_i)^{-1} d_i \sigma_{i1} &= \\ \sigma_{i1}^{ii} Id_i \sigma_{i1} &= \sigma_{i1}^{ii} d_i \sigma_{i1}. \end{aligned}$$

Durch Prämultiplikation des so erhaltenen ersten und dritten Summanden mit obigem Q_{i1} bekommt man

$$(R - (n_i + r_i) - \alpha_{i1} - 1) (Z_i' Z_i)^{-1} d_i \sigma_{i1}.$$

Dies ist genau Nagars Resultat (vgl. dessen Theorem 1).

In ähnlicher Weise läßt sich nachweisen, daß auch Nagars Resultat über die Momentenmatrix der k -Klasse ein Spezialfall von (4.8) ist. Auf eine genaue Herleitung wird hier verzichtet.

5. Beweis von Theorem 4.1 und Theorem 4.2

Der Beweis erfolgt entsprechend dem in Abschnitt 3 beschriebenen Vorgehen für die k -Klasse: Zunächst wird der Stichprobenfehler $e(K) = \hat{\alpha}(K) - \alpha$ der KMC-Schätzfunktionen (4.2) ausgedrückt als Funktion der Vorherbestimmten, der wahren (red.-Form)-Parameter und der Störglieder:

$$(5.1) \quad e(K) = \hat{\alpha}(K) - \alpha = (I + QF)^{-1} Q \left(\sum_{h=1}^N \left(s^{ih} (Z_i' + (1 - k_{ih}) V_i' + k_{ih} V_i^{N*}) \right) u_h \right)_i$$

$$\text{mit } F = \left(s^{ih} (V_i' Z_h + Z_i' V_h + (1 - k_{ih}) V_i' V_h + k_{ih} V_i^{N*} V_h) \right)_{i,h}.$$

Für $(I + QF)^{-1}$ wird dann entsprechend (3.3) die Entwicklung

$$(5.2) \quad (I + QF)^{-1} = \sum_{s=0}^{\infty} (-1)^s (QF)^s \quad \text{zu Hilfe genommen.}$$

Als erstes wird gezeigt, daß die Ausdrücke (5.1) und (4.2) tatsächlich identisch sind. Wegen $y_h = \bar{y}_h \alpha_h + u_h$ läßt sich (4.2) auch schreiben als

$$(5.3) \quad e(K) = \hat{\alpha}(K) - \alpha = \left(s^{ih} (\bar{z}_i' \bar{z}_h - k_{ih} \hat{v}_i' \hat{v}_h) \right)^{-1} \left(\sum_{h=1}^N s^{ih} (\bar{z}_i' - k_{ih} \hat{v}_i') u_h \right)_{i,h}$$

Die Inverse auf der rechten Seite von (5.3) ist gleich dem Ausdruck $(I + QF)^{-1} Q$ auf der rechten Seite von (5.1), denn

$$(I + QF)^{-1} Q = (Q^{-1} + F)^{-1} = \left(s^{ih} (Z_i' Z_h + V_i' Z_h + Z_i' V_h + (1 - k_{ih}) V_i' V_h + k_{ih} V_i^{N*} V_h) \right)^{-1}_{i,h}$$

$$\begin{aligned}
 &= \left(s^{ih} \left[\begin{pmatrix} \pi_i^1 X' X \pi_h & \pi_i^1 X' X_h \\ X_j^1 X \pi_h & X_j^1 X_h \end{pmatrix} + \begin{pmatrix} \bar{v}_i^1 X \pi_h & \bar{v}_i^1 X_h \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} \pi_i^1 X' \bar{v}_h & 0 \\ X_j^1 \bar{v}_h & 0 \end{pmatrix} + \begin{pmatrix} (1-k_{ih}) \bar{v}_i^1 \bar{v}_h & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} k_{ih} \bar{v}_i^1 X (X' X)^{-1} X' \bar{v}_h & 0 \\ 0 & 0 \end{pmatrix} \right]_{i,h}^{-1} \right) \\
 &= \left(s^{ih} \begin{pmatrix} \pi_i^1 X' X \pi_h + \bar{v}_i^1 X \pi_h + \pi_i^1 X \bar{v}_h + \bar{v}_i^1 \bar{v}_h - k_{ih} \hat{v}_i^1 \hat{v}_h & \pi_i^1 X' X_h + \bar{v}_i^1 X_h \\ X_j^1 X \pi_h + X_j^1 \bar{v}_h & X_j^1 X_h \end{pmatrix} \right)_{i,h}^{-1} \\
 &= \left(s^{ih} \begin{pmatrix} v_i^1 v_h - k_{ih} \hat{v}_i^1 \hat{v}_h & v_i^1 X_h \\ X_j^1 v_h & X_j^1 X_h \end{pmatrix} \right)_{i,h}^{-1} \\
 &= \left(s^{ih} (\bar{z}_i^1 \bar{z}_h - k_{ih} \hat{v}_i^1 \hat{v}_h) \right)_{i,h}^{-1} \quad \text{q.e.d.}
 \end{aligned}$$

Wenno zeigt man, daß der zweite Ausdruck auf der rechten Seite von (5.3), nämlich

$$\left(\sum_h s^{ih} (\bar{z}_i^1 - k_{ih} \hat{v}_i^1) u_h \right)_i$$

gleich ist dem Ausdruck

$$\left(\sum_h s^{ih} (\bar{z}_i^1 + (1-k_{ih}) v_i^1 + k_{ih} v_i^1 N^*) u_h \right)_i$$

auf der rechten Seite von (5.1). Es ist nämlich:

$$\begin{aligned}
 &\left(\sum_h s^{ih} (\bar{z}_i^1 + (1-k_{ih}) v_i^1 + k_{ih} v_i^1 N^*) u_h \right)_i = \\
 &= \left(\sum_h s^{ih} \begin{pmatrix} \pi_i^1 X + (1-k_{ih}) \bar{v}_i^1 + k_{ih} \bar{v}_i^1 X (X' X)^{-1} X' \\ X_j^1 \end{pmatrix} u_h \right)_i =
 \end{aligned}$$

$$\begin{aligned}
 &= \left(\sum_h s^{jh} \begin{pmatrix} y_i' - k_{jh} \hat{v}_i' \\ x_i' \end{pmatrix} v_h \right)_i = \\
 &= \left(\sum_h s^{jh} (\bar{z}_i' - k_{jh} \hat{v}_i') v_h \right)_i \quad \text{q.e.d.}
 \end{aligned}$$

Damit ist die Gleichheit von (5.1) mit (5.3) und damit auch mit (4.2) gezeigt.

5.1 Die Berechnung des Bias

Zur Berechnung des (approximativen) Klein-Stichproben-Bias wird nun die bereits erwähnte Potenzreihenentwicklung von $(I + QP)^{-1}$ zu Hilfe genommen (siehe (5.2)). Es ist

$$(5.4) \quad (I + QP)^{-1} = \sum_{s=0}^{\infty} (-1)^s (QP)^s \stackrel{\text{appr.}}{=} I - QP,$$

wobei appr. bedeutet, daß nur die ersten beiden Glieder der unendlichen Reihe in (5.4) berücksichtigt werden. Da in (5.1)

$$Q = O(T^{-1}) \text{ und } \left(\sum_h s^{jh} (z_i' + (1 - k_{jh}) v_i' + k_{jh} v_i^{N*}) v_h \right)_i = O(T^{1/2})$$

und außerdem $QP = O(T^{-1/2})$ würden höhere Potenzen in QP insgesamt nur einen Beitrag liefern, der kleiner als $O(T^{-1})$ ist. Daher kann man sich auf die ersten beiden Glieder der Potenzreihe beschränken, wenn man bis zur $O(T^{-1})$ exakte Ergebnisse wünscht.

Unter Berücksichtigung der Approximation (5.4) ist nun (5.1) gleich (appr.)

$$(5.5) \quad \alpha(K) - \hat{\alpha}(V) - \alpha - (I - QPQ) \left(\sum_h s^{jh} (z_i' + (1 - k_{jh}) v_i' + k_{jh} v_i^{N*}) v_h \right)_i.$$

Unter Beachtung von (4.4), der Voraussetzung über die k_{jh} , treten nun allerdings auch in (5.5) noch Terme auf, die kleiner als $O(T^{-1})$ sind und die man bei der hier gewünschten Approximation vernachlässigen kann. So ist z.B. in (5.5)

$$Q\left(\sum_h s^{ih} k_{ih} v_i^* u_h\right)_i = Q\left(\sum_h s^{ih} v_i^* u_h\right)_i + Q\left(\sum_h s^{ih} T^{-1} x_{ih} v_i^* u_h\right)_i,$$

wobei $T^{-1} x_{ih} v_i^* u_h = O(T^{-1})$ und somit der zweite Summand auf der rechten Seite der obigen Gleichung wegen $Q = O(T^{-1})$ gleich $O(T^{-2})$, also vernachlässigbar klein ist. Vernachlässigt man also alle in (5.5) auftretende Terme mit kleinerer Wachstumsordnung als T^{-1} , so bekommt man folgenden Ausdruck für den Stichprobenfehler der KMC-Schätzfunktion:

$$\begin{aligned} (5.6) \quad e(K) &= \hat{\alpha}(K) - \alpha = \\ &= Q\left[\left(\sum_h s^{ih} z_i^* u_h\right)_i - T^{-1} \left(\sum_h s^{ih} x_{ih} v_i^* u_h\right)_i + \left(\sum_h s^{ih} v_i^* u_h\right)_i\right] \\ &\quad - Q\left(s^{ih} z_i^* v_h\right)_{i,h} Q\left(\sum_h s^{ih} z_i^* u_h\right)_i \\ &\quad - Q\left(s^{ih} v_i^* z_h\right)_{i,h} Q\left(\sum_h s^{ih} z_i^* u_h\right)_i \end{aligned}$$

Von der rechten Seite von (5.6) ist jetzt der Erwartungswert zu bilden, was wegen der bekannten Eigenschaften des Erwartungswertoperators gliedweise, d.h. summandenweise, geschehen kann. Es ist also, da Q nichtstochastisch ist:

$$E\left\{Q\left(\sum_h s^{ih} z_i^* u_h\right)_i\right\} = Q\left(\sum_h s^{ih} z_i^* E\{u_h\}\right)_i = 0,$$

da z_i nichtstochastisch und $E\{u_h\}$ nach Voraussetzung gleich 0 ist.

$$E\left\{T^{-1} Q\left(\sum_h s^{ih} x_{ih} v_i^* u_h\right)_i\right\} = T^{-1} Q\left(\sum_h s^{ih} x_{ih} E\{v_i^* u_h\}\right)_i =$$

$$(5.7) \quad Q\left(\sum_h s^{ih} x_{ih} q_{ih}\right)_i, \text{ da}$$

$$E\{v_i^* u_h\} = D_i^* E\{u_h\} = D_i^* T \sigma_h = T q_{ih}$$

$$E\left\{Q\left(\sum_h s^{ih} v_i^* u_h\right)_i\right\} = Q\left(\sum_h s^{ih} E\{v_i^* u_h\}\right)_i =$$

$$(5.8) = Q \left(\sum_h s^{ih} R q_{ih} \right)_i, \quad \text{da}$$

$$\begin{aligned} E\{V_i' N^* u_h\} &= D_i' E\{U' X (X' X)^{-1} X' u_h\} = D_i' \left(E\{u_k' N^* u_h\} \right)_k = \\ &= D_i' \left(E\{\text{tr}(u_k' N^* u_h)\} \right)_k = D_i' \left(\text{tr}(N^* E\{u_h u_k'\}) \right)_k = \\ &= D_i' \left(\sigma_{hk} \text{tr}((X' X)^{-1} (X' X)) \right)_k = D_i' \sigma_h^R = q_{ih}^R, \text{ wenn} \end{aligned}$$

man noch berücksichtigt, daß $E\{u_h u_k'\} = \begin{pmatrix} \sigma_{hk} & 0 \\ 0 & \sigma_{hk} \end{pmatrix}$.

Damit sind die Erwartungswerte der ersten drei Summanden auf der rechten Seite von (5.6) berechnet.

Als nächstes ist der Erwartungswertoperator auf den vierten Summanden in (5.6) anzuwenden:

$$\begin{aligned} Q E\{ & (s^{ih} z_{ij}' v_h)_{i,h} (q_{hj})_{h,j} \left(\sum_k s^{ik} z_{ij}' u_k \right)_j = \\ & - Q E\{ \left(\sum_{h,k,j} s^{ih} s^{ik} z_{ij}' v_h q_{hj} z_{ij}' u_k \right)_i \} , \text{ wenn man berücksichtigt,} \end{aligned}$$

daß für partitionierte Matrizen dieselben Multiplikationsregeln wie für gewöhnliche Matrizen gelten, sofern die Dimensionen der Untermatrizen zueinander passen. **Obiger Ausdruck ist**

$$- Q \left(\sum_{h,k,j} s^{ih} s^{ik} z_{ij}' v_h q_{hj} z_{ij}' u_k \right)_i \quad \text{wegen der elementweisen}$$

und summandenweisen Anwendbarkeit des Erwartungswertoperators.

$$= Q \left(\sum_{h,k,j} s^{ih} s^{ik} z_{ij}' z_{ij} q_{jh} D_h' \sigma_k \right)_i \quad \text{wegen (A1)!}$$

$$(5.9) = Q \left(\sum_{h,k} s^{ih} \sum_j s^{ik} z_{ij}' z_{ij} q_{jh} q_{hk} \right)_i .$$

Wendet man schließlich den Erwartungswertoperator auf den fünften Summanden in (5.6) an, so ergibt sich:

1) vgl. Anhang A2.

$$\begin{aligned}
 & \mathbb{Q}E\left\{ (s^{ih}v_i'z_h)_{i,h} (\varrho_{hj})_{h,j} \left(\sum_k s^{ik}z_j'u_k \right)_j \right\} = \\
 & \mathbb{Q}E\left\{ \left(\sum_{h,k,j} s^{ih}s^{ik}v_i'z_h\varrho_{hj}z_j'u_k \right)_i \right\} = \\
 & \mathbb{Q}\left(\sum_{h,k,j} s^{ih}s^{ik}E\{v_i'z_h\varrho_{hj}z_j'u_k\} \right)_i = \\
 & \mathbb{Q}\left(\sum_{h,k,j} s^{ih}s^{ik}\text{tr}(z_h\varrho_{hj}z_j')q_{ik} \right)_i \quad \text{wegen (A2).}
 \end{aligned}$$

Es ist aber

$$\begin{aligned}
 & \sum_j s^{jk}\text{tr}(z_h\varrho_{hj}z_j') = \text{tr}\left(\sum_j s^{jk}\varrho_{hj}z_j'z_h\right) = \text{tr}(I) = \\
 & = n_h + r_h, \quad \text{wenn } h=k \quad \left((s^{ih}z_i'z_h)^{-1} = (\varrho_{ih})_{i,h} \right).
 \end{aligned}$$

$n_h + r_h$ ist die Anzahl der auf der rechten Seite der h -ten Gleichung vorkommenden erklärenden Variablen.

Der obige Ausdruck läßt sich also zerlegen in

$$(5.10) \quad \mathbb{Q}\left(\sum_h s^{ih}(n_h + r_h)q_{ih}\right)_i + \mathbb{Q}\left(\sum_{\substack{h,k \\ h \neq k}} s^{ih}s^{ik}\text{tr}\left(\sum_j \varrho_{hj}z_j'z_h\right)q_{ik}\right)_i$$

Zusammenfassung der vier Ausdrücke (5.7), (5.8), (5.9) und (5.10) ergibt genau den in Theorem 4.1 angegebenen approximativen Klein-Stichproben-Bias der KMC-Schätzfunktion.

Damit ist Theorem 4.1 bewiesen.

5.2 Die Berechnung der Momentenmatrix

Im Gegensatz zu der noch relativ einfachen Herleitung des Bias zur $O(T^{-1})$ benötigt man zur Herleitung der Momentenmatrix zur $O(T^{-2})$ doch schon einen beträchtlich größeren Rechenaufwand, da ja hier der quadratische Ausdruck $\alpha(K)e(K)' = (\hat{\alpha}(K) - \alpha)(\hat{\alpha}(K) - \alpha)'$ Grundlage der Berechnung ist und man außerdem noch die Glieder der Ordnung T^{-2} berücksichtigen muß. (Die Glieder der Ordnung T^{-1} bilden gerade die bekannte asymptotische Kovarianzmatrix).

Wegen (5.1) ist

$$(5.11) \quad e(K)e(K)' = (I + QP)^{-1} QGQ(I + QP)'^{-1}, \text{ wo } P \text{ schon in (5.1)}$$

definiert ist und

$$G := \left(\sum_h s^{ih} (z_i' + (1 - k_{ih})v_i' + k_{ih}v_i'^* u_h) \right)_i \times \left(\sum_h s^{ih} (z_i' + (1 - k_{ih})v_i' + k_{ih}v_i'^* u_h) \right)_i'$$

Es ist hier zu beachten, daß sowohl G als auch P wegen der Symmetrie von $\Sigma^{-1} = (s^{ih})_{i,h}$ und $K = (k_{ih})_{i,h} = K'$ (vgl. (4.4)) beides symmetrische Matrizen sind.

Nimmt man nun wieder für $(I + QP)^{-1}$ die Potenzreihenentwicklung (5.2) zu Hilfe und verwendet diesmal zur Approximation die ersten drei Glieder (die höheren Potenzen sind von kleinerer Wachstumsordnung als T^{-2}), so erhält man:

$$(I + QP)^{-1} = \sum_{s=0}^{\infty} (-1)^s (QP)^s \approx I - QP + QPQ$$

Einsetzen dieses Ausdrucks in (5.11) und Ausmultiplizieren ergibt wegen der Symmetrie von P und G :

$$(I - QP + QPQ)QGQ(I - PQ + PQPQ) =$$

$$= \underline{QGG} - \underline{QGQFQ} + \underline{QGQFQFQ} - \underline{QFQGG} + \underline{QFQGGFQ} - \underline{QFQGGQFQFQ} \\ + \underline{QFQFQGG} - \underline{QFQFQGGFQ} + \underline{QFQFQGGQFQFQ}$$

Da $G = O(T)$, $F = O(T^{1/2})$ und $Q = O(T^{-1})$, sind die unterstrichenen Glieder von kleinerer Wachstumsordnung als T^{-2} und können daher vernachlässigt werden.

Prä- und post-multipliziert man noch (5.11) mit Q^{-1} , so ergibt sich schließlich approximativ:

$$(5.12) \quad Q^{-1}e(K)e(K)'Q^{-1} = G - (GQF + FQG) + \\ + (GQFQF + FQGQF + FQFQG)$$

In der weiteren Berechnung können daher auf der rechten Seite von (5.12) Glieder von kleinerer Wachstumsordnung als $O(1)$ vernachlässigt werden, da $Q = O(T^{-1})$.

Auf der rechten Seite von (5.12) kann jetzt gliedweise der Erwartungswert berechnet werden. Dazu werden G und F zerlegt in

$$G = G_1 + G_{1/2} + G_0 \quad \text{bzw.}$$

$$F = F_{1/2} + F_0,$$

wobei G_1 den Teil von G bezeichnet, der $O(T^1)$ ist, $G_{1/2}$ den Teil von G , der $O(T^{1/2})$ ist, und G_0 den Teil von G , der $O(T^0) = O(1)$ ist. Entsprechendes gilt für F .

Für G , das in (5.11) als Produkt von Spalten- und Zeilenvektor definiert ist, ergibt sich ausmultipliziert:

$$G = \left(\sum_h s^{ih} (z_i^{*} + v_i^{*} n_{ih}^{*} - (1/T) \alpha_{ih} v_i^{*} (I - N^{*}) u_h) \right)_{i,1} \times \\ \left(\sum_k s^{jk} v_k^{*} (z_j^{*} + n_{jk}^{*} v_j^{*} - (1/T) \alpha_{jk} (I - N^{*}) v_j^{*}) \right)_{1,j}, \quad \text{da} \\ k_{ih} = 1 + \frac{1}{T} \alpha_{ih} \\ = \left(\sum_h s^{ih} (z_i^{*} + v_i^{*} n_{ih}^{*} + (1/T) \alpha_{ih} v_i^{*} n_{ih}^{*}) u_h \right)_{i,1} \times \\ \left(\sum_k s^{jk} v_k^{*} (z_j^{*} + n_{jk}^{*} v_j^{*} + (1/T) \alpha_{jk} n_{jk}^{*} v_j^{*}) \right)_{1,j},$$

wobei $N_{ih} = N'_{ih} := -(1/T) \alpha_{ih} I + N^* \quad (= O(T^{-1}))$

$$= \left(\sum_{h,k} s^{ih} s^{jk} (Z'_{ih} u'_{jk} Z'_{ij} + Z'_{ih} u'_{jk} N'_{jk} V_{ij} + V_{ij} N'_{ih} u'_{jk} Z'_{ij} + \right. \\ \left. + V_{ij} N'_{ih} u'_{jk} N'_{jk} V_{ij}) \right)_{i,j} + \text{Glieder von } O(T^{-1/2})$$

und kleiner, die vernachlässigt werden können.

Es ist daher:

$$G_1 = \left(\sum_{h,k} s^{ih} s^{jk} Z'_{ih} u'_{jk} Z'_{ij} \right)_{i,j} \\ G_{1/2} = \left(\sum_{h,k} s^{ih} s^{jk} Z'_{ih} u'_{jk} N'_{jk} V_{ij} \right)_{i,j} + \\ (5.13) \quad + \left(\sum_{h,k} s^{ih} s^{jk} V_{ij} N'_{ih} u'_{jk} Z'_{ij} \right)_{i,j}$$

$$G_0 = \left(\sum_{h,k} s^{ih} s^{jk} V_{ij} N'_{ih} u'_{jk} N'_{jk} V_{ij} \right)_{i,j}$$

$$\text{mit } N_{ih} = -(1/T) \alpha_{ih} I + X(X'X)^{-1} X'$$

Ebenso erhält man aus der Definition von F in (5.1):

$$F_{1/2} = (s^{ih} y'_{ij} Z'_{jh})_{i,h} + (s^{ih} Z'_{ij} V_{jh})_{i,h} \\ (5.14) \\ F_0 = (s^{ih} V_{ij} N'_{ih} V_{jh})_{i,h}$$

Entsprechend Gleichung (5.12) wird als erstes der Erwartungswert von G , das in (5.11) definiert ist, berechnet, wobei auf die Zerlegung in (5.13), $G = G_1 + G_{1/2} + G_0$ zurückgegriffen wird.

Es ist

$$E(G_1) = \left(\sum_{h,k} s^{ih} s^{jk} E\{Z'_{ih} u'_{jk} Z'_{ij}\} \right)_{i,j} \\ = \left(\sum_{h,k} s^{ih} s^{jk} \sigma_{hk} Z'_{ij} Z'_{ij} \right)_{i,j}, \text{ da } E\{u'_{jk} u'_{jk}\} = \sigma_{hk} I$$

$$\begin{aligned}
 E(G_{1/2}) &= \left(\sum_{h,k} s^{ih} s^{jk} z_i^1 E \{ u_h u_k^{N_{jk} V_{ij}} \} \right)_{i,j} + \dots \\
 &= 0, \text{ da } E \{ u_h u_k^{N_{jk} V_{ij}} \} = E \{ u_h u_k^{N_{jk} \Pi_{ij}} \} D_{ij} = 0, \text{ da die} \\
 &\quad \text{Störlieder zeitlich unabhängig normalverteilt} \\
 &\quad \text{angenommen wurden. Dasselbe Argument gilt für} \\
 &\quad \text{den zweiten Summanden von } G_{1/2} \text{ in (5.13).}
 \end{aligned}$$

$$\begin{aligned}
 E(G_0) &= \left(\sum_{h,k} s^{ih} s^{jk} w_{ij}^1 E \{ u_h u_k^{N_{jk} V_{ij}} \} \right)_{i,j} = \\
 &= \left(\sum_{h,k} s^{ih} s^{jk} p_i^1 E \{ \Pi_{ij}^1 u_h u_k^{N_{jk} \Pi_{ij}} \} \right)_{i,j} = \\
 &= \left(\sum_{h,k} s^{ih} s^{jk} p_i^1 [R(\sigma_k \sigma_h^1 + \sigma_{hk} \Sigma) + (R - \alpha_{ih})(R - \alpha_{jk}) \sigma_h \sigma_k^1] D_{ij} \right)_{i,j} \\
 &\quad \text{wegen (B3)} \\
 &= \left(\sum_{h,k} s^{ih} s^{jk} (R q_{ik} q_{jh}^1 + (R - \alpha_{ih})(R - \alpha_{jk}) q_{ih} q_{jk}^1 + R \sigma_{hk} \Omega_{ij}) \right)_{i,j}
 \end{aligned}$$

Insgesamt ist also

$$\begin{aligned}
 (5.15) \quad E(G) &= \left(\sum_{h,k} s^{ih} s^{jk} (\sigma_{hk} z_i^1 z_j^1 + R q_{ik} q_{jh}^1 + (R - \alpha_{ih})(R - \alpha_{jk}) q_{ih} q_{jk}^1 + \right. \\
 &\quad \left. + R \sigma_{hk} \Omega_{ij}) \right)_{i,j}
 \end{aligned}$$

Damit ist der erste Summand auf der rechten Seite von (5.12) erledigt.

Als nächstes ist der Erwartungswert des zweiten Summanden in (5.12) zu berechnen. Es ist unter Verwendung der Zerlegung in (5.13) und (5.14):

$$\begin{aligned}
 E(G \mathcal{P}) &= E \{ (G_1 + G_{1/2} + G_0) \mathcal{Q}(\mathcal{P}_{1/2} + \mathcal{P}_0) \} \\
 &= E \{ G_1 \mathcal{Q} \mathcal{P}_{1/2} + (G_1 \mathcal{Q} \mathcal{P}_0 + G_{1/2} \mathcal{Q} \mathcal{P}_{1/2}) \} + \text{Glieder kleinerer} \\
 &\quad \text{Wachstumsordnung als } O(1), \text{ wenn man beachtet, daß} \\
 &\quad \mathcal{Q} = O(T^{-1}).
 \end{aligned}$$

Nun ist

$$E \{ G_1 \mathcal{Q} \mathcal{P}_{1/2} \} =$$

$$\begin{aligned}
 & - \mathbb{E} \left\{ \left(\sum_{h,k} s^{ih} s^{ik} z_i^u u_k^v z_j \right)_{i,j} (\gamma_{j1})_{j,1} \left[(s^{lm} v_l^z z_m)_{1,m} + (s^{lm} z_l^v v_m)_{1,m} \right] \right\} \\
 & - \mathbb{E} \left\{ \left(\sum_{h,k,j,1} s^{ih} s^{ik} z_i^u u_k^v z_j \gamma_{j1} s^{lm} u_l^v z_m \right)_{i,m} \right\} + \\
 & + \mathbb{E} \left\{ \left(\sum_{h,k,j,1} s^{ih} s^{ik} z_i^u u_k^v z_j \gamma_{j1} s^{lm} z_l^u v_m \right)_{i,m} \right\}
 \end{aligned}$$

= 0, da die Störliefer zeitlich unabhängig normalverteilt sind.

$$\mathbb{E} \{ G_1 O P_0 \} =$$

$$\begin{aligned}
 & = \mathbb{E} \left\{ \left(\sum_{h,k} s^{ih} s^{ik} z_i^u u_k^v z_j \right)_{i,j} (\gamma_{j1})_{j,1} (s^{lm} v_l^u v_m^v)_{1,m} \right\} \\
 & = \mathbb{E} \left\{ \left(\sum_{h,k,j,1} s^{ih} s^{ik} s^{lm} z_i^u u_k^v z_j \gamma_{j1} v_l^u v_m^v \right)_{i,m} \right\} \\
 & = \left(\sum_{h,k,j,1} s^{ih} s^{ik} s^{lm} z_i^u u_k^v z_j \gamma_{j1} v_l^u v_m^v \right)_{i,m} \\
 & = \left(\sum_{h,k,j,1} s^{ih} s^{ik} s^{lm} z_i^u u_k^v z_j \gamma_{j1} (a_{1h} a_{mk}^u + a_{1k} a_{mh}^u + (R_{1m}^{-\infty}) \sigma_{hk} \Omega_{1m}) \right)_{i,m} \\
 & \quad \text{nach (B4).}
 \end{aligned}$$

Schließlich ist unter Beachtung von (5.13) bzw. (5.14) und den üblichen Multiplikationsregeln für partitionierte Matrizen:

$$\mathbb{E} \{ G_{1/2} P_{1/2} \} =$$

$$\begin{aligned}
 & \mathbb{E} \left\{ \left(\sum_{h,k,j,1} s^{ih} s^{ik} s^{lm} z_i^u u_k^v z_j \gamma_{j1} v_l^u v_m^v \right)_{i,m} \right\} + \\
 & + \mathbb{E} \left\{ \left(\sum_{h,k,j,1} s^{ih} s^{ik} s^{lm} z_i^u u_k^v z_j \gamma_{j1} z_l^v v_m \right)_{i,m} \right\} + \\
 & + \mathbb{E} \left\{ \left(\sum_{h,k,j,1} s^{ih} s^{ik} s^{lm} v_l^u u_k^v z_j \gamma_{j1} z_l^v v_m \right)_{i,m} \right\} + \\
 & + \mathbb{E} \left\{ \left(\sum_{h,k,j,1} s^{ih} s^{ik} s^{lm} v_l^u u_k^v z_j \gamma_{j1} z_l^v v_m \right)_{i,m} \right\}
 \end{aligned} \tag{5.16}$$

Der erste Term in (5.16) ergibt:

$$\begin{aligned} & \left(\sum_{h,k,j,l} s^{ih} s^{jk} s^{lm} Z_i^1 R \xi_{ih}^{u,u} u_{jk}^{m,m} D_{jl}^{m,m} D_{jl}^{m,m} Z_m^1 \right)_{i,m} = \\ & = \left(\sum_{h,k,j,l} s^{ih} s^{jk} s^{lm} Z_i^1 Z_m^1 [(R - \alpha_{jk}) q_{jl}^1 q_{jl}^1 q_{jh}^1 \sigma_{hk} + \text{tr}(q_{jl} \Sigma_{1,j}) + q_{jh}^1 q_{jl}^1 q_{lk}^1] \right)_{i,m} \\ & \hspace{15em} \text{nach (B5).} \end{aligned}$$

Der zweite Term in (5.16) ist

$$\begin{aligned} & \left(\sum_{h,k,j,l} s^{ih} s^{jk} s^{lm} Z_i^1 R \xi_{ih}^{u,u} u_{jk}^{m,m} D_{jl}^{m,m} Z_l^1 U^1 D_m^1 \right)_{i,m} = \\ & = \left(\sum_{h,k,j,l} s^{ih} s^{jk} s^{lm} Z_i^1 Z_l^1 q_{lj} [\sigma_{hk} \Sigma_{jm} + q_{jh} q_{mk}^1 + (R - \alpha_{jk}) q_{jk} q_{mh}^1] \right)_{i,m} \\ & \hspace{15em} \text{nach (B6).} \end{aligned}$$

Für den dritten Term in (5.16) erhält man

$$\begin{aligned} & \left(\sum_{h,k,j,l} s^{ih} s^{jk} s^{lm} D_l^1 R \xi_{ih}^{u,u} u_{jk}^{m,m} Z_j^1 D_{jl}^{m,m} Z_m^1 \right)_{i,m} = \\ & = \left(\sum_{h,k,j,l} s^{ih} s^{jk} s^{lm} [\sigma_{hk} \Sigma_{jl} + q_{lk} q_{jh}^1 + (R - \alpha_{jh}) q_{ih} q_{jl}^1] q_{lj} Z_l^1 Z_m^1 \right)_{i,m} \\ & \hspace{15em} \text{nach (B7).} \end{aligned}$$

Der vierte Term in (5.16) ergibt schließlich

$$\begin{aligned} & \left(\sum_{h,k,j,l} s^{ih} s^{jk} s^{lm} D_l^1 R \xi_{ih}^{u,u} u_{jk}^{m,m} Z_j^1 Z_l^1 U^1 D_m^1 \right)_{i,m} = \\ & = \left(\sum_{h,k,j,l} s^{ih} s^{jk} s^{lm} \text{tr}(Z_j^1 q_{jl}^1 Z_l^1) [q_{il} q_{mh}^1 + (R - \alpha_{jh}) q_{ih} q_{ml}^1 + \sigma_{hl} \Sigma_{im}] \right)_{i,m} \\ & \hspace{15em} \text{nach (B8).} \end{aligned}$$

Setzt man nun diese vier Ausdrücke mit dem bereits berechneten Ausdruck für $R \xi_{ij}^{G, Q} P_0$ zusammen, so bekommt man für $R \xi_{ij}^{G, Q} P_0$ folgenden Ausdruck:

$$R \xi_{ij}^{G, Q} P_0 = -R(\alpha) =$$

(5.17)

$$\begin{aligned}
 &= \left(\sum_{h,k,j,l} s^{ih} s^{jk} s^{lm} z_i' z_j' z_l' (q_{lh} q_{mk}' + q_{lk} q_{mh}' + (R - \alpha_{lm}) \sigma_{hk} \Omega_{lm}) \right)_{i,m} \\
 &+ \left(\sum_{h,k,j,l} s^{ih} s^{jk} s^{lm} z_i' z_m' [(R - \alpha_{jk}) q_{jl}' z_j' q_{lh} + \sigma_{hk} \text{tr}(Q_{jl} \Omega_{lj}) + q_{jh}' z_{jl} q_{lk}'] \right)_{i,m} \\
 &+ \left(\sum_{h,k,j,l} s^{ih} s^{jk} s^{lm} z_i' z_l' z_j' [\sigma_{hk} \Omega_{jm} + q_{jh} q_{mk}' + (R - \alpha_{jk}) q_{jk} q_{mh}'] \right)_{i,m} \\
 &+ \left(\sum_{h,k,j,l} s^{ih} s^{jk} s^{lm} [\sigma_{hk} \Omega_{il} + q_{il} q_{lh}' + (R - \alpha_{ih}) q_{ih} q_{lk}'] z_l' z_j' z_m' \right)_{i,m} \\
 &+ \left(\sum_{h,k,j,l} s^{ih} s^{jk} s^{lm} \text{tr}(z_j' z_{jl} z_l') [q_{il} q_{mh}' + (R - \alpha_{ih}) q_{ih} q_{mk}' + \sigma_{hk} \Omega_{im}'] \right)_{i,m}
 \end{aligned}$$

Damit ist der zweite Summand in (5.12) berechnet. Der dritte Summand in (5.12) ist

$$FQG = F'Q'G' = (G)F',$$

wegen der Symmetrie von F, G und Q, also gerade die transponierte Matrix des zweiten Summanden. Damit ist aber auch der Erwartungswert von FQG die Transponierte der Matrix in (5.17), also

$$(5.18) \quad E\{FQG\} = -B(\mathcal{L})'$$

Als nächstes ist der Erwartungswert des vierten Summanden in (5.12) zu berechnen, nämlich der Erwartungswert von

$$\begin{aligned}
 GQFQF &= (G_1 + G_{1/2} + G_0) Q (F_{1/2} + F_0) Q (F_{1/2} + F_0) \\
 &= G_1 Q F_{1/2} Q F_{1/2} + \text{Glieder kleinerer Wachstumsordnung} \\
 &\quad \text{als } O(1), \text{ die ja bei unserer Approximation nicht interessieren.}
 \end{aligned}$$

Dies ergibt sich, wenn man die Zerlegung von G und F in (5.13) bzw. (5.14) sowie die Wachstumsordnungen der einzelnen Matrizen beachtet.

Es ist aber nach (5.13) und (5.14)

$$(5.19) \quad G_1 \mathcal{Q}^{\mathbb{P}} 1/2 \mathcal{Q}^{\mathbb{P}} 1/2 =$$

$$\begin{aligned} & \left(\sum_{h,k} s^{ih} s^{jk} Z_i^u u_h^v Z_j^z \right)_{i,j} (\mathcal{Q}_{jl})_{j,l} (s^{lm} Z_l^v Z_m^u)_{l,m} (\mathcal{Q}_{mn})_{m,n} (s^{nq} Z_n^v Z_q^u)_{n,q} \\ & + \left(\sum_{h,k} s^{ih} s^{jk} Z_i^u u_h^v Z_j^z \right)_{i,j} (\mathcal{Q}_{jl})_{j,l} (s^{lm} Z_l^v Z_m^u)_{l,m} (\mathcal{Q}_{mn})_{m,n} (s^{nq} Z_n^v Z_q^u)_{n,q} \\ & + \left(\sum_{h,k} s^{ih} s^{jk} Z_i^u u_h^v Z_j^z \right)_{i,j} (\mathcal{Q}_{jl})_{j,l} (s^{lm} Z_l^v Z_m^u)_{l,m} (\mathcal{Q}_{mn})_{m,n} (s^{nq} Z_n^v Z_q^u)_{n,q} \\ & + \left(\sum_{h,k} s^{ih} s^{jk} Z_i^u u_h^v Z_j^z \right)_{i,j} (\mathcal{Q}_{jl})_{j,l} (s^{lm} Z_l^v Z_m^u)_{l,m} (\mathcal{Q}_{mn})_{m,n} (s^{nq} Z_n^v Z_q^u)_{n,q} \end{aligned}$$

Der Erwartungswert des ersten Terms in (5.19) ist

$$\begin{aligned} & \left(\sum_{h,k,j,l,m,n} s^{ih} s^{jk} s^{lm} s^{nq} Z_i^u u_h^v Z_j^z \mathcal{E} \{ u_h^v Z_j^z \mathcal{Q}_{jl}^{\mathbb{P}} \mathcal{Q}_{lm}^{\mathbb{P}} \mathcal{Q}_{mn}^{\mathbb{P}} \mathcal{Q}_{nq}^{\mathbb{P}} \} \right)_{i,q} = \\ & \left(\sum_{h,k,j,l,m,n} s^{ih} s^{jk} s^{lm} s^{nq} Z_i^u (\sigma_{hk} Z_j^z \mathcal{Q}_{jl}^{\mathbb{P}} \mathcal{Q}_{ln}^{\mathbb{P}} \mathcal{Q}_{nm}^{\mathbb{P}} Z_m^u + Z_m^u \mathcal{Q}_{mn}^{\mathbb{P}} \mathcal{Q}_{nq}^{\mathbb{P}} \mathcal{Q}_{jl}^{\mathbb{P}} Z_j^z + \right. \\ & \quad \left. + \mathcal{Q}_{nh}^{\mathbb{P}} \mathcal{Q}_{nm}^{\mathbb{P}} Z_m^u Z_j^z \mathcal{Q}_{jl}^{\mathbb{P}} \mathcal{Q}_{lq}^{\mathbb{P}}) Z_q^u \right)_{i,q} \end{aligned}$$

nach (B9),

der Erwartungswert des zweiten Terms in (5.19)

$$\begin{aligned} & \left(\sum_{h,k,j,l,m,n} s^{ih} s^{jk} s^{lm} s^{nq} Z_i^u \mathcal{E} \{ u_h^v Z_j^z \mathcal{Q}_{jl}^{\mathbb{P}} \mathcal{Q}_{ln}^{\mathbb{P}} Z_m^u \mathcal{Q}_{mn}^{\mathbb{P}} \mathcal{Q}_{nq}^{\mathbb{P}} \} \right)_{i,q} = \\ & \left(\sum_{h,k,j,l,m,n} s^{ih} s^{jk} s^{lm} s^{nq} Z_i^u [\sigma_{hk} Z_j^z \mathcal{Q}_{jl}^{\mathbb{P}} \mathcal{Q}_{ln}^{\mathbb{P}} \text{tr}(Z_m^u \mathcal{Q}_{mn}^{\mathbb{P}} Z_n^u) + Z_m^u \mathcal{Q}_{mn}^{\mathbb{P}} Z_n^u Z_j^z \mathcal{Q}_{jl}^{\mathbb{P}} \mathcal{Q}_{lq}^{\mathbb{P}} \mathcal{Q}_{nq}^{\mathbb{P}} + \right. \\ & \quad \left. + Z_n^u \mathcal{Q}_{nm}^{\mathbb{P}} Z_m^u Z_j^z \mathcal{Q}_{jl}^{\mathbb{P}} \mathcal{Q}_{lq}^{\mathbb{P}} \mathcal{Q}_{nq}^{\mathbb{P}}] \right)_{i,q} \\ & \text{nach (B10),} \end{aligned}$$

Es muß zwischen dem Index q und q_{lh} , das als $q_{lh} = \mathcal{D}_i^{\mathbb{P}} \sigma_h$ erklärt ist, unterschieden werden. Aus dem Zusammenhang heraus dürfte jeweils klar sein, welches q gemeint ist.

den Erwartungswert des dritten Terms in (5.19)

$$\left(\sum_{h,k,j,l,m,n} s^{ih} s^{jk} s^{lm} s^{pq} Z_i^* E_{hk}^{\dagger} Z_j^* Z_l^* Z_m^* Z_n^* Z_q \right)_{i,q} =$$

$$\left(\sum_{h,k,j,l,m,n} s^{ih} s^{jk} s^{lm} s^{pq} Z_i^* E_{hk}^{\dagger} Z_j^* Z_l^* \text{tr}(\Omega_{mn} \Omega_{nm}) + Z_l^* Z_j^* Z_i^* a_{mh}' \Omega_{mn} \Omega_{nk} + \right.$$

$$\left. + \text{tr}(Z_j^* Z_l^* Z_i^*) a_{mk}' \Omega_{mn} \Omega_{ph} \right)_{i,q}$$

nach (B11),

den Erwartungswert des vierten Terms in (5.19)

$$\left(\sum_{h,k,j,l,m,n} s^{ih} s^{jk} s^{lm} s^{pq} Z_i^* E_{hk}^{\dagger} Z_j^* Z_l^* Z_m^* Z_n^* Z_q \right)_{i,q} =$$

$$\left(\sum_{h,k,j,l,m,n} s^{ih} s^{jk} s^{lm} s^{pq} Z_i^* E_{hk}^{\dagger} Z_j^* Z_l^* Z_n^* \Omega_{nm} \Omega_{mq} + Z_l^* Z_j^* Z_i^* Z_n^* \Omega_{nm} \Omega_{mh} a_{qk}' + \right.$$

$$\left. + \text{tr}(Z_j^* Z_l^* Z_i^*) Z_n^* \Omega_{nm} \Omega_{mk} a_{qh}' \right)_{i,q}$$

nach (B12).

Zusammenfassung der eben berechneten vier Erwartungswerte ergibt:

$$(5.20) \quad E\{Q^2 P^2 P^2\} = G =$$

$$\left(\sum_{h,k,j,l,m,n} s^{ih} s^{jk} s^{lm} s^{pq} Z_i^* (E_{hk}^{\dagger} Z_j^* Z_l^* \Omega_{ln} \Omega_{nm} Z_m' + Z_m' \Omega_{mn} \Omega_{nk} a_{lh}' Z_j^* Z_l^* + \right.$$

$$\left. + a_{nh}' \Omega_{nm} Z_m' Z_j^* Z_l^* a_{lk}^{\dagger}) Z_q \right)_{i,q} +$$

$$\left(\sum_{h,k,j,l,m,n} s^{ih} s^{jk} s^{lm} s^{pq} Z_i^* E_{hk}^{\dagger} Z_j^* Z_l^* \Omega_{ln} \text{tr}(\Omega_{mn} \Omega_{nm}) + Z_m' \Omega_{mn} Z_n^* Z_j^* Z_l^* a_{lh} a_{qk}' + \right.$$

$$\left. + Z_n^* \Omega_{nm} Z_m' Z_j^* Z_l^* a_{lk} a_{qh}' \right)_{i,q} +$$

$$\left(\sum_{h,k,j,l,m,n} s^{ih} s^{jk} s^{lm} s^{pq} Z_i^* E_{hk}^{\dagger} Z_j^* Z_l^* \text{tr}(\Omega_{mn} \Omega_{nm}) + Z_l^* Z_j^* Z_i^* a_{mh}' \Omega_{mn} \Omega_{nk} + \right.$$

$$\left. + \text{tr}(Z_j^* Z_l^* Z_i^*) a_{mk}' \Omega_{mn} \Omega_{nh} \right)_{i,q} +$$

(noch 5.20)

$$\left(\sum_{h,k,i,l,m,o} s^{ih} s^{jk} s^{lm} s^{pq} Z_i^l \sigma_{hk} Z_j^l Z_{jl}^l Z_o^l \rho_{pm} \Omega_{nq} + Z_l Z_{lj} Z_j^l Z_o^l \rho_{om} \sigma_{mh} \sigma_{qk}' + \right. \\ \left. + \text{tr}(Z_j Z_{jl} Z_l^l) Z_o^l \rho_{om} \sigma_{mk} \sigma_{qh}' \right)_{i,q}$$

Damit sind die ersten vier Summanden von (5.12) erledigt.

Der Erwartungswert des fünften Summanden auf der rechten Seite von (5.12) ist unter Vernachlässigung der Glieder kleinerer Wachstumsordnungen als $O(1)$ und unter Beachtung der Zerlegungen (5.13), (5.14) und der Wachstumsordnungen der einzelnen Matrizen gleich

$$\begin{aligned} E\{PQGQ^P\} &= E\{(P_{1/2} + P_0) Q(G_1 + G_{1/2} + G_0) Q(P_{1/2} + P_0)\} \\ &= E\{P_{1/2} Q G_1 Q^P_{1/2}\} \text{ zur } O(1) \end{aligned}$$

Wegen (5.13) und (5.14) ist nun aber

$$(5.21) \quad P_{1/2} Q G_1 Q^P_{1/2} =$$

$$\begin{aligned} &(s^{il} v_i^l Z_l)_{i,l} (\gamma_{lj})_{l,j} \left(\sum_{h,k} s^{jh} s^{mk} Z_j^l u_{h'k'} Z_m \right)_{j,m} (\gamma_{mo})_{m,o} (s^{pq} v_o^p Z_q)_{o,q} \\ &+ (s^{il} v_i^l Z_l)_{i,l} (\gamma_{lj})_{l,j} \left(\sum_{h,k} s^{jh} s^{mk} Z_j^l u_{h'k'} Z_m \right)_{j,m} (\gamma_{mo})_{m,o} (s^{pq} Z_o^p v_q)_{o,q} \\ &+ (s^{il} Z_i^l v_l)_{i,l} (\gamma_{lj})_{l,j} \left(\sum_{h,k} s^{jh} s^{mk} Z_j^l u_{h'k'} Z_m \right)_{j,m} (\gamma_{mo})_{m,o} (s^{pq} v_o^p Z_q)_{o,q} \\ &+ (s^{il} Z_i^l v_l)_{i,l} (\gamma_{lj})_{l,j} \left(\sum_{h,k} s^{jh} s^{mk} Z_j^l u_{h'k'} Z_m \right)_{j,m} (\gamma_{mo})_{m,o} (s^{pq} Z_o^p v_q)_{o,q} \end{aligned}$$

Der Erwartungswert des ersten Terms in (5.21) ist

$$\begin{aligned} &\left(\sum_{h,k,l,j,m,o} s^{il} s^{jh} s^{mk} s^{pq} D_i^l E_j^p U^l Z_l \gamma_{lj} Z_j^l u_{h'k'} Z_m \gamma_{mo} v_o^p U^q Z_q \right)_{i,q} = \\ &\left(\sum_{h,k,l,j,m,o} s^{il} s^{jh} s^{mk} s^{pq} [a_{ih} a_{o'k'} \gamma_{om} Z_m^l + \text{tr}(Z_{lj} Z_l^l Z_j^l) + a_{il} a_{o'h} \gamma_{om} Z_m^l Z_l^l \gamma_{lj} Z_j^l] \right. \\ &\quad \left. + \sigma_{hk} \Omega_{lo} \gamma_{om} Z_m^l Z_j^l \gamma_{lj} Z_l^l Z_q \right)_{i,q} \end{aligned}$$

nach (B13),

der Erwartungswert des zweiten Terms in (5.21)

$$\left(\sum_{h,k,l,j,i,m,n} s^{il} s^{jh} s^{mk} s^{nq} D_i^{\dagger} B \{ \Pi_l^{\dagger} \Omega_{l,j} Z_{j,i}^{\dagger} u_{h,k}^{\dagger} Z_m \Omega_{mn} Z_n^{\dagger} \Pi_q \} D_q \right)_{i,q} =$$

$$\left(\sum_{h,k,l,j,i,m,n} s^{il} s^{jh} s^{mk} s^{nq} [\alpha_{ih} \alpha_{ql}^{\dagger} \text{tr}(\Omega_{l,j} Z_{j,i}^{\dagger} Z_l) \text{tr}(\Omega_{mn} Z_n^{\dagger} Z_m) + \right.$$

$$\left. + \alpha_{ik} \alpha_{qh}^{\dagger} \text{tr}(Z_l \Omega_{l,j} Z_{j,i}^{\dagger} Z_p \Omega_{nm} Z_m^{\dagger}) + \sigma_{hk} \Omega_{iq} \text{tr}(Z_l \Omega_{l,j} Z_{j,i}^{\dagger} Z_m \Omega_{mn} Z_n^{\dagger})] \right)_{i,q}$$

nach (B14),

der Erwartungswert des dritten Terms in (5.21)

$$\left(\sum_{h,k,l,j,i,m,n} s^{il} s^{jh} s^{mk} s^{nq} Z_i^{\dagger} B \{ \Pi_l^{\dagger} \Omega_{l,j} Z_{j,i}^{\dagger} u_{h,k}^{\dagger} Z_m \Omega_{mn} Z_n^{\dagger} \} D_q \right)_{i,q} =$$

$$\left(\sum_{h,k,l,j,i,m,n} s^{il} s^{jh} s^{mk} s^{nq} Z_i^{\dagger} [Z_j \Omega_{j,i} \alpha_{lh} \alpha_{nk}^{\dagger} \Omega_{nm} Z_m^{\dagger} + Z_m \Omega_{mn} \alpha_{nh} \alpha_{lk}^{\dagger} \Omega_{l,j} Z_j^{\dagger} + \right.$$

$$\left. + \sigma_{hk} \text{tr}(Z_m \Omega_{mn} \Omega_{nl} \Omega_{l,j} Z_j^{\dagger})] Z_q \right)_{i,q}$$

nach (B15),

der Erwartungswert des vierten Terms in (5.21)

$$\left(\sum_{h,k,l,j,i,m,n} s^{il} s^{jh} s^{mk} s^{nq} Z_i^{\dagger} B \{ \Pi_l^{\dagger} \Omega_{l,j} Z_{j,i}^{\dagger} u_{h,k}^{\dagger} Z_m \Omega_{mn} Z_n^{\dagger} \} D_q \right)_{i,q} =$$

$$\left(\sum_{h,k,l,j,i,m,n} s^{il} s^{jh} s^{mk} s^{nq} Z_i^{\dagger} [Z_j \Omega_{j,i} \alpha_{lh} \alpha_{nk}^{\dagger} \text{tr}(Z_m \Omega_{mn} Z_n^{\dagger}) + \right.$$

$$\left. + Z_m \Omega_{mn} Z_n^{\dagger} \Omega_{j,i} \alpha_{lk} \alpha_{qh}^{\dagger} + \sigma_{hk} Z_n \Omega_{nm} Z_m^{\dagger} \Omega_{j,i} \Omega_{l,q}] \right)_{i,q}$$

nach (B16).

Wast man nun die so berechneten vier Erwartungswerte wieder zusammen, so erhält man

$$B \{ \Pi \Omega \Omega^{\dagger} \} = D =$$

(5.22)

$$\begin{aligned}
 &= \left(\sum_{h,k,l,j,m,n} s^{il} s^{jh} s^{mk} s^{nq} [q_{ih} q'_{nk} \varrho_{nm} z'_m \text{tr}(\varrho_{lj} z'_j z'_l) + q_{ih} q'_{nh} \varrho_{nm} z'_m z'_l \varrho_{lj} z'_j + \right. \\
 &\quad \left. + \sigma_{hk} \varrho_{ln} \varrho_{nm} z'_m z'_j \varrho_{jl} z'_j z'_q] \right)_{i,q} + \\
 &\left(\sum_{h,k,l,j,m,n} s^{il} s^{jh} s^{mk} s^{nq} [q_{ih} q'_{nk} \text{tr}(\varrho_{lj} z'_j z'_l) \text{tr}(\varrho_{mn} z'_m z'_n) + \right. \\
 &\quad \left. + q_{ik} q'_{qh} \text{tr}(\varrho_{lj} \varrho_{lj} z'_j z'_n \varrho_{nm} z'_m) + \sigma_{hk} \varrho_{iq} \text{tr}(\varrho_{lj} \varrho_{lj} z'_j z'_m \varrho_{mn} z'_n) \right] \Big)_{i,q} + \\
 &\left(\sum_{h,k,l,j,m,n} s^{il} s^{jh} s^{mk} s^{nq} z'_i [z'_j \varrho_{jl} q_{lh} q'_{nk} \varrho_{nm} z'_m + z'_m \varrho_{mn} q_{nh} q'_{lk} \varrho_{lj} z'_j + \right. \\
 &\quad \left. + \sigma_{hk} \text{tr}(\varrho_{mn} \varrho_{nl} z'_j z'_j) \right] z'_q \Big)_{i,q} + \\
 &\left(\sum_{h,k,l,j,m,n} s^{il} s^{jh} s^{mk} s^{nq} z'_j [z'_j \varrho_{jl} q_{lh} q'_{nk} \text{tr}(\varrho_{mn} z'_m z'_n) + \right. \\
 &\quad \left. + z'_m \varrho_{mn} z'_j z'_j \varrho_{jl} q_{lk} q'_{qh} + \sigma_{hk} z'_n \varrho_{nm} z'_m z'_j \varrho_{jl} \varrho_{lq} \right] \Big)_{i,q}
 \end{aligned}$$

Damit sind bis auf den letzten alle Summanden in (5.12) erledigt. Dieser letzte Summand ist jedoch

$$PQFQG = P'Q'P'Q'G' = (C)PQF',$$

also wegen der Symmetrie von F, G und Q gerade die Transponierte des vierten Summanden von (5.12). Also ist der Erwartungswert

$$(5.23) \quad E\{PQFQG\} = G', \text{ wo } G \text{ in (5.20) definiert ist.}$$

Faßt man schließlich die Ausdrücke (5.15), (5.17), (5.18), (5.20), (5.22) und (5.23) zusammen und prä- und post-multipliziert wieder mit (dem nichtstochastischen) Q , so bekommt man genau das Ergebnis von Theorem 4.2 für die KMC-Momentenmatrix zur $O(T^{-2})$.

), E. D.

Es ist noch folgendes zu beachten:

Faßt man in der oben angegebenen Weise die Ausdrücke zusammen, so bleibt als einziger Term der $O(T^{-1})$ die Matrix (s. 5.15)

$$(5.24) \quad Q \left(\sum_{h,k} s^{ih} s^{jk} \sigma_{hk} z_i' z_j \right)_{i,j} Q$$

Wegen $s = \sigma$ (Annahme (4.3)) und der Symmetrie von Σ ist die mittlere Matrix in (5.24) gleich

$$\left(\sum_{h,k} \sigma^{ih} \sigma_{hk} \sigma^{kj} z_i' z_j \right)_{i,j} = \left(z_i' z_j \sum_{h,k} \sigma^{ih} \sigma_{hk} \sigma^{kj} \right)_{i,j}$$

Nun ist aber

$$\begin{aligned} \sum_{h,k} \sigma^{ih} \sigma_{hk} \sigma^{kj} &= \sum_h \sigma^{ih} \sum_k \sigma_{hk} \sigma^{kj} = \sigma^{ij}, \text{ da} \\ \sum_k \sigma_{hk} \sigma^{kj} &= \begin{cases} 1 & \text{für } h=j \\ 0 & \text{sonst.} \end{cases} \end{aligned}$$

Insgesamt ist (5.24) also gleich

$$Q(\sigma^{ij} z_i' z_j) Q = Q Q^{-1} Q = Q$$

d.h., wie nicht anders zu erwarten, die bekannte asymptotische 3SLS-Kovarianzmatrix, da ja $\text{plim} \sqrt{T}(K-E) = \text{plim}(1/\sqrt{T})X = 0$.

Schlußbemerkungen, Zusammenfassung der Ergebnisse

Die Hauptergebnisse und -aspekte der vorliegenden Arbeit sollen hier kurz zusammengefaßt werden:

Ausgehend vom allgemeinen linearen interdependenten ökonomischen Modell unter den üblichen Annahmen wurde in Kapitel I kurz das Problem der Identifikation beleuchtet: In unserem Fall ist Identifizierbarkeit gleichbedeutend mit einer eindeutigen Beziehung zwischen Strukturform und reduzierter Form des Modells.

In Kapitel II wurden alle Schätzverfahren, die üblicherweise in interdependenten Systemen verwendet werden, unter verschiedenen Gesichtspunkten entwickelt. Dabei ergaben sich einige interessante neue Aspekte: Während schon lange bekannt ist, daß alle Limited-Information-Schätzverfahren zu einer Klasse, der k -Klasse gehören, zeigt sich, daß sich auch alle Full-Information-Schätzverfahren in einer Klasse, der sogenannten K -Matrix-Klasse, zusammenfassen lassen. Auch h -Klasse und Doppel- k -Klasse haben jeweils ein Full-Information-Pendant, nämlich die h -Vektor- und die Doppel- K -Matrix-Klasse. Insgesamt zeigte sich, daß die Schätzverfahren der k -Klasse die Spezialfälle der Schätzverfahren der K -Matrix-Klasse sind, bei denen in der Schätzung die kontemporären Kovarianzen von verschiedenen Gleichungen des Systems nicht berücksichtigt werden. Diese Vernachlässigung hat eine größere asymptotische Kovarianzmatrix der II-Schätzverfahren zur Folge, sofern man das gesamte System betrachtet. Bedingungen für asymptotische Schätzeigenschaften der KMC-Schätzfunktionen erhielt man in Analogie zu den Bedingungen bei den k -Klasse-Schätzverfahren.

Folgende weitere neue Gesichtspunkte ergaben sich im

Verläufe von Kapitel II: Alle LI- und FI-Schätzverfahren haben eine Interpretation als Instrumenten-Variablen-Schätzfunktion¹⁾. Definiert man jedoch in Analogie zum verallgemeinerten Kleinst-Quadrate-(GLS-) Schätzverfahren eine verallgemeinerte Instrumenten-Variablen-(GIV-) Schätzfunktion, so erkennt man, daß man alle FI-Schätzverfahren als GIV-Schätzverfahren betrachten kann.

Als weitere Interpretationsmöglichkeit der einzelnen Schätzverfahren bot sich die Darstellung der Schätzverfahren als Ergebnis von Minimierungsprozessen an. Dabei minimieren die LI-Schätzverfahren jeweils denselben Ausdruck wie ihr FI-Pendant, allerdings unter Verwendung verschiedener "Maße": Während bei den LI-Verfahren die Spur als Maß zur Verwendung kommt, ist es bei den FI-Verfahren die Determinante²⁾. Hierdurch kommt auch zum Ausdruck, daß die FI-Verfahren mehr Information als die LI-Verfahren verwerten, denn die Spur berücksichtigt ja nur die Hauptdiagonale einer Matrix, während die Determinante alle Elemente berücksichtigt.

Kapitel II macht also in besonderem Maße deutlich, daß sich auch unter verschiedenen Interpretationsgesichtspunkten alle LI- und FI-Schätzverfahren bei interdependenten Modellen analytisch einheitlich behandeln lassen. In diesem Zusammenhang sei auf Abschnitt 3.1 hingewiesen, in dem die GLS-Schätzung des bekannten Zellner-Modells der scheinbar unverbundenen Regressionen übertragen auf ein interdependentes System als FI-Verallgemeinerung der Gleichungs-OLS-Schätzung interpretiert wurde. Daß die 3SLS-Schätzfunktion die FI-Verallgemeinerung der 2SLS-

1) vgl. Tabelle 1b) auf S. 121

2) vgl. Tabelle 1a) auf S. 121

Schätzfunktion ist, erkennt man unmittelbar aus den Schätzgleichungen. Daß jedoch auch FIML zu LIML im selben Verhältnis wie 3SLS zu 2SLS steht, kam meines Erachtens in der Literatur bisher noch nicht genügend klar zum Ausdruck. In Abschnitt 3.3. wurde dies durch eine einfache Erweiterung der Log-Likelihood-Funktion deutlich. Daß sich die FIML-Schätzwerte nicht im selben Maße analytisch darstellen lassen wie die LIML-Schätzwerte, liegt daran, wie Abschnitt 3.3 zeigte, daß man die Log-Likelihood-Funktion bei FIML nicht ähnlich weit konzentrieren kann wie bei LIML, sondern schon frühzeitig auf ein nichtlineares Gleichungssystem stößt, was auf der Verwendung der Kovarianzen der kontemporären Störglieder verschiedener Gleichungen bei FIML beruht.

Schließlich wurde in Abschnitt 5.1 eine bekannte Beziehung zwischen 2SLS-Schätzfunktion und den Doppel-k-Klasse-Schätzfunktionen auf die entsprechende Beziehung zwischen 3SLS und einer Unterklasse der Doppel-K-Matrix-Klasse verallgemeinert. Es zeigte sich, daß auch diese Unterklasse (die als Doppel-h-Vektor-Klasse bezeichnet wurde) als FI-Verallgemeinerung der Doppel-k-Klasse betrachtet werden kann. Letztlich wurden in Abschnitt 5.2 die Zusammenhänge zwischen den verschiedenen Schätzverfahren in einigen Extremfällen betrachtet.

In Kapitel III wurde das eigentliche Hauptergebnis der vorliegenden Arbeit entwickelt: Ein Verfahren von Nagar zur Bestimmung von Bias zur $O(T^{-1})$ und Momentenmatrix zur $O(T^{-2})$ der k-Klasse-Schätzfunktionen wurde auf die Schätzfunktionen der K-Matrix-Klasse verallgemeinert. Die in Theorem 4.1 und Theorem 4.2 enthaltenen Ergebnisse für die K-Matrix-Klasse bestätigen Nagars Ergebnisse für die k-Klasse, da sie diese als Spezialfall enthalten.

Die vorliegende Arbeit wurde in der Absicht geschrie-

ben, die theoretisch-analytischen Zusammenhänge zwischen den verschiedenen Schätzverfahren deutlich zu machen, und vor allen Dingen zu zeigen, daß sich (fast) alle Konzepte der Einzelgleichungs-II-Schätzung auch auf die Simultan-FI-Schätzung übertragen lassen, so daß man zu einer einheitlichen Schätztheorie bei linearen interdependenten Systemen gelangt. Dieser rein theoretischen Ausrichtung entsprechend wurde auf numerische Beispiele zu den einzelnen Schätzverfahren verzichtet. Bei den in der Praxis vorkommenden oft sehr großen ökonometrischen Modellen treten jedoch bei Anwendung der Schätzverfahren teilweise beträchtliche Schwierigkeiten auf. Dies liegt vor allem an zwei Punkten: Zum einen fehlt es bei großen Modellen an genügend Beobachtungsmaterial, um die Schätzverfahren in ihrer ursprünglichen Form anwenden zu können (vgl. II,5.2.). Zum anderen treten oft Komplikationen bei der Programmierung der einzelnen Schätzmethoden auf: Besonders bei den Vorschätzungen benutzenden ausführbaren FI-Schätzverfahren werden wegen der Kompliziertheit der Schätzgleichungen die Computerprogramme lang und unübersichtlich und man stößt schnell an die Grenze der Speicherplatzkapazität der Rechenanlagen. Wegen dieser Schwierigkeiten wird auf größere Modelle meist nur die 2SLS-Methode (als einfachste die Interdependenz berücksichtigende Methode) angewandt. Ein Beispiel für die Anwendung der 3SLS-Methode auf ein größeres Modell findet man bei Friedrich¹⁾, der eine Version des Vierteljahresmodells für die Bundesrepublik Deutschland von D.Lüdeke²⁾ auch 3SLS-schätzte (mit Prinzipalkomponenten).

1) Friedrich, D.: a.a.O. S.90f.

2) Lüdeke, D.: Ein ökonometrisches Vierteljahresmodell für die Bundesrepublik Deutschland, Tübingen 1969.

Die oben genannten Schwierigkeiten sind jedoch nicht prinzipieller Natur: Durch Erhebungen von ökonomischen Größen in kürzeren Zeitabständen könnte man den Umfang des Beobachtungsmaterials vergrößern und so durch diese zusätzliche Information die Güte und Anwendungsmöglichkeiten der Schätzverfahren positiv beeinflussen. Die Fortschritte in der Entwicklung der Computerindustrie lassen erwarten, daß auch die Programmierschwierigkeiten und Speicherplatzbeschränkungen mit der Zeit an Bedeutung verlieren. In diesem Zusammenhang wäre eine speziell auf die Bedürfnisse der Ökonometrie zugeschnittene Programmiersprache wünschenswert. Sobald also diese Schwierigkeiten behoben sind, kann die größere Effizienz der FI-Schätzverfahren auch in der Praxis bei größeren Modellen voll zum tragen kommen.

Weitere Schwierigkeiten bei der Anwendung der Schätzverfahren auf ökonometrische Modelle ergeben sich aus ökonomischen Überlegungen: Selbst wenn man eine lineare Beziehung zwischen verschiedenen ökonomischen Größen nur als Approximation betrachtet, wird sie sich in manchen Fällen nicht über den gesamten Beobachtungszeitraum aufrecht erhalten lassen, und es können sogenannte Strukturbrüche auftreten. Insbesondere die konjunkturellen Schwankungen innerhalb des Beobachtungszeitraums müßten besser berücksichtigt werden. Ein wichtiger neuer Schritt in diese Richtung wurde in dem Artikel von D. Lüdeke¹⁾ gemacht, in dem eine Investitionsfunktion nach konjunkturellen Aufschwungs- und Abschwungsphasen jeweils getrennt geschätzt wird, was zu wesentlich besseren Schätzergebnissen führt.

1) Lüdeke, D.: Strukturveränderungen des Investitionsverhaltens im Konjunkturverlauf, Jahrbücher für Nationalökonomie und Statistik, Bd.188, 1973, S.289-319.

Auch läßt sich die Annahme der Nicht-Autokorrelation der Störglieder in der Praxis oft nicht aufrecht erhalten. Ob sich eine ähnlich einheitliche Schätztheorie auch bei Vorliegen von Autokorrelation entwickeln läßt, ist noch eine offene Frage und ein interessantes Gebiet weiterer theoretischer Forschung. Zwar führt Hendry¹⁾ eine Klassifikation²⁾ von Schätzverfahren auch bei Berücksichtigung von Autokorrelation durch, doch handelt es sich dabei nicht um eine Klassifikation in dem Sinne, daß man (wie bei der K-Matrix-Klasse) durch Variation eines oder mehrerer Parameter alle Schätzverfahren erhalten würde. Bei Hendrys Klassifikation müssen vielmehr ganze Matrizen ausgetauscht werden.

1) Hendry, D.F.: The Structure of Simultaneous Equations Estimators, *Journal of Econometrics* 4, 1976, 51-88.

2) Im Zusammenhang der Klassifizierung von Schätzverfahren sei auch auf den Artikel von Keller, W.J.: A New Class of Limited-Information Estimators for Simultaneous Equation Systems, *Journal of Econometrics* 3, 1975, 71-92, hingewiesen, der sich allerdings nur auf LI-Verfahren bezieht.

ANHANG

A1: Konventionen, Notationen

Die Kapitel dieser Arbeit sind mit römischen Ziffern durchnummeriert (I,II,III). Innerhalb eines Kapitels sind die Abschnitte und Unterabschnitte jeweils mit arabischen Ziffern nummeriert (z.B. 2.1., 2.3.1.), Gleichungen sind innerhalb eines Abschnitts durch Ziffern, die mit runden Klammern versehen sind, fortlaufend nummeriert (z.B. (3.12)). Wird innerhalb eines Kapitels auf eine Gleichung oder einen Abschnitt verwiesen, so geschieht dies, indem die Gleichungsnummer oder die Nummer des (Unter-) Abschnitts direkt angegeben wird; wird auf eine Gleichung oder einen Abschnitt außerhalb des laufenden Kapitels verwiesen, so wird zusätzlich zur Gleichungs- oder Abschnittsnummer noch die Kapitelnummer vorgesetzt (also z.B. I,(1.1) oder II, 3.2.1.).

Die in dieser Arbeit verwendeten analytischen Verfahren und Begriffe beschränken sich auf relativ elementare Grundlagen der mathematischen Statistik und Wahrscheinlichkeitsrechnung sowie der linearen Algebra. Begriffe wie z.B. Erwartungswert, Dichtefunktion, Zufallsvariable oder Matrix, Vektor, Eigenwertproblem werden als bekannt vorausgesetzt. Ihre ausführliche Erklärung findet man zu Beginn oder im Anhang von fast allen ökonometrischen Standardwerken, ebenso die Darstellung von Matrix- bzw. Vektordifferentiation. Auf eine nochmalige Darstellung wird daher hier verzichtet. Es sei jedoch kurz erwähnt, daß mit

|| die Determinante einer Matrix,

tr die Spur einer Matrix (=die Summe der Hauptdiagonalelemente),

rg der Rang einer Matrix (=die Maximalzahl der linear unabhängigen Zeilen- bzw. Spaltenvektoren)

bezeichnet wird.

Es wurde der Übersichtlichkeit wegen versucht, mit möglichst wenigen Symbolen auszukommen. Dies bedingte, daß oft verschiedene Begriffe mit demselben Symbol belegt wurden. So kann z.B. Y eine Zufallsvariable oder eine Beobachtungsmatrix bezeichnen, α einen "wahren" Parameter oder eine Variable, U eine Matrix von Störgliedern oder eine Funktion von α . Aus dem Zusammenhang heraus ist jedoch jeweils klar, was gemeint ist. Das Symbol $\hat{\cdot}$ bezeichnet eine Schätzfunktion.

Besondere Schwierigkeiten macht die einfache und doch durchschaubare Darstellung von partitionierten Matrizen, die besonders für die Darstellung der FI-Methoden notwendig sind. Meist wird in der Literatur dabei auf das sogenannte Kronecker-Produkt zweier Matrizen zurückgegriffen: Unter dem Kronecker-Produkt einer $M \times N$ -Matrix A und einer $R \times T$ -Matrix B versteht man die $MR \times NT$ -Matrix

$$A \otimes B = \begin{pmatrix} a_{11}^B & a_{12}^B & \dots & a_{1N}^B \\ a_{21}^B & \dots & & \vdots \\ \vdots & & & \vdots \\ a_{M1}^B & \dots & \dots & a_{MN}^B \end{pmatrix}, \text{ wobei } A = (a_{ik}).$$

Die Darstellung von FI-Schätzfunktionen mit Hilfe des Kronecker-Produkts verlangt jedoch meist umfangreiche Definitionen von Matrizen. Die Darstellung ist dann zwar kompakt und elegant, man erkennt aus ihr jedoch nicht unmittelbar die Bedeutung der einzelnen Matrizen. Bei der Definition der K-Matrix-Klasse scheint die Darstellung in Kronecker-Produkt-Schreibweise nicht möglich zu sein (zumindest nicht ohne komplizierte "Vor"-Definitionen von Matrizen). In der vorliegenden Arbeit fand daher meist eine andere Darstellungsweise Verwendung, die auch unmittelbar die Struktur der partitionierten Matrizen erkennen läßt:

Es seien A_{ik} ($i=1, \dots, M; k=1, \dots, N$) beliebige Matrizen,

die sich zu der partitionierten Matrix

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1N} \\ A_{21} & \dots & & \vdots \\ \vdots & & & \vdots \\ A_{M1} & \dots & \dots & A_{MN} \end{pmatrix}$$

zusammensetzen lassen sollen. In Kurzschreibweise läßt sich A dann schreiben als

$$A = (A_{ik})_{\substack{i=1,\dots,M \\ k=1,\dots,N}} \text{ oder } (A_{ik})_{i,k}, \text{ wenn klar ist, wieviel}$$

"Spalten" bzw. "Zeilen" von Untermatrizen gemeint sind.

Ist A_{ik} z.B. von der speziellen Form $B_i B_k$, so kann man A schreiben als

$$A = (B_i B_k)_{i,k}.$$

$$A = (A_{i1})_{i,1} \text{ bedeutet } \begin{pmatrix} A_{11} \\ A_{21} \\ \vdots \\ A_{M1} \end{pmatrix} \text{ und wird auch oft abge-}$$

kürzt durch $(A_{i1})_i$. A_{ik} kann natürlich auch als Spezialfall einen Untervektor darstellen.

$A = (A_{ii})_{i,i}$ bezeichnet die blockdiagonale Matrix

$$\begin{pmatrix} A_{11} & & 0 \\ & A_{22} & \\ 0 & & \ddots \\ & & & A_{NN} \end{pmatrix}$$

Das Kronecker-Produkt $A \otimes B$ läßt sich in dieser Schreibweise darstellen als

$$A \otimes B = (a_{ik} B)_{i,k}.$$

Da die Rechenregeln für partitionierte Matrizen wie die von einfachen Matrizen sind, sofern die Zeilen und Spalten der Untermatrizen jeweils zusammenpassen, gilt z.B.:

$$AB = (A_{ik})_{i,k} (B_{jk})_{j,k} = (\sum_j A_{ij} B_{jk})_{i,k}.$$

Nachdem die Darstellungsweise von partitionierten Matrizen erläutert wurde, hier noch eine Erklärung des Begriffes der Wachstumsordnung:

Das Symbol $O(\)$ bezeichnet sowohl die stochastische als auch die nichtstochastische Wachstumsordnung:

1. Es sei x_T ein von T abhängiger nichtstochastischer Skalar, $f(T)$ eine Funktion von T . Dann bedeutet

$$x_T = O(f(T)),$$

daß $|x_T/f(T)| < c$ mit einer Konstanten c für alle T .

2. Es sei x_T ein von T abhängiger stochastischer Skalar. Dann bedeutet

$$x_T = O(f(T)),$$

daß zu jedem $\varepsilon > 0$ ein $z(\varepsilon)$ existiert, so daß

$$P\{|x_T/f(T)| > z(\varepsilon)\} < \varepsilon \quad \text{für alle } T.$$

Für Matrizen bedeutet $O(\)$, daß alle Matricelemente die obigen Bedingungen erfüllen.

Sei z.B. X eine nichtstochastische $T \times R$ -Matrix, deren Elemente dem Absolutbetrag nach beschränkt sind (durch c) für alle T . Dann ist $X'X = O(T)$, denn für das (i,k) -te Element $\tilde{x}_{ik}$ von $X'X$ gilt wegen der Beschränktheit der Elemente von X folgende Abschätzung:

$$|\tilde{x}_{ik}| \leq \sum_{t=1}^T |x_{it}| |x_{tk}| < \sum_{t=1}^T c^2 = Tc^2, \text{ so daß}$$

$$|\tilde{x}_{ik}/T| < c^2,$$

also nach obiger Definition $X'X$ tatsächlich $O(T)$ ist.

Zum Begriff der Konsistenz und anderer in dieser Arbeit vorkommenden Schätzeigenschaften wird auf die ökonometrischen Standardwerke, die auf Seite 5 und 6 zitiert sind, verwiesen.

A2: Die Berechnung der Erwartungswerte

Zur Berechnung des Bias und der Momentenmatrix werden die folgenden Erwartungswerte benötigt:

A. Zur Berechnung des Bias

$$(A1) \ E\{V_{jh}Z_{jh}^i u_k\} = Z_{jh}D_{jh}'\sigma_k$$

$$(A2) \ E\{V_{jh}^i Z_{jh} Z_{jh}^i u_k\} = \text{tr}(Z_{jh}Z_{jh}^i)D_{jh}'\sigma_k$$

B. Zur Berechnung der Momentenmatrix

$$(B3) \ E\{u_i' u_h u_k' u_j\} = R(\sigma_k \sigma_h' + \sigma_{hk} \Sigma) + (R - x_{jh})(R - x_{jk})\sigma_h \sigma_k'$$

$$(B4) \ E\{u_i u_k' Z_{jh} D_{jh} u_l u_m\} = Z_{jh} D_{jh} D_{jh}' (\sigma_h \sigma_k' + \sigma_k \sigma_h' + (R - x_{lm})\sigma_{hl} \Sigma)$$

$$(B5) \ E\{u_i u_k' u_j u_l' D_{jh} D_{jh}' u_l u_m\} = (R - x_{jk})\sigma_k' D_{jh} D_{jh}' \sigma_h I + \sigma_{hk} \text{tr}(D_{jh} D_{jh}' u_l u_m \Sigma) N'' + \sigma_h' D_{jh} D_{jh}' D_{jh}' \sigma_k N''$$

$$(B6) \ E\{u_i u_k' u_j u_l' D_{jh} D_{jh}' Z_{jh} u_l\} = (R - x_{jk}) Z_{jh} D_{jh}' \sigma_k \sigma_h' + \sigma_{hk} Z_{jh} D_{jh}' \sigma_h \Sigma + Z_{jh} D_{jh}' D_{jh}' \sigma_h \sigma_k'$$

$$(B7) \ E\{u_i u_k' u_j u_l' Z_{jh} D_{jh}' u_l u_m\} = (R - x_{jh})\sigma_h \sigma_k' D_{jh} D_{jh}' Z_{jh} + \sigma_{hk} \Sigma D_{jh} D_{jh}' Z_{jh} + \sigma_k \sigma_h' D_{jh} D_{jh}' Z_{jh}$$

$$(B8) \ E\{u_i u_k' u_j u_l' Z_{jh} Z_{jh}' u_l u_m\} = \text{tr}(Z_{jh} Z_{jh}' u_l u_m) ((R - x_{jh})\sigma_h \sigma_k' + \sigma_{hk} \Sigma + \sigma_k \sigma_h')$$

$$(B9) \ E\{u_i u_k' Z_{jh} D_{jh}' u_l u_m u_n u_p\} = \sigma_{hk} Z_{jh} D_{jh}' D_{jh}' \Sigma D_{jh} D_{jh}' \sigma_n' + Z_{jh} D_{jh}' D_{jh}' \sigma_k \sigma_h' D_{jh} D_{jh}' Z_{jh}' + \sigma_h' D_{jh} D_{jh}' \sigma_n' Z_{jh} D_{jh}' D_{jh}' \sigma_k I$$

$$(B10) \ E\{u_i u_k' Z_{jh} D_{jh}' u_l u_m u_n u_p\} = \sigma_{hk} Z_{jh} D_{jh}' D_{jh}' \Sigma \text{tr}(Z_{jh} D_{jh}' \sigma_n') + Z_{jh} D_{jh}' D_{jh}' \sigma_n' Z_{jh} D_{jh}' D_{jh}' \sigma_k' + \sigma_h \sigma_k' + Z_{jh} D_{jh}' \sigma_n' Z_{jh} D_{jh}' D_{jh}' \sigma_k \sigma_h'$$

$$(B11) \ E\{u_i u_k' Z_{jh} D_{jh}' u_l u_m u_n u_p\} = \sigma_{hk} Z_{jh} D_{jh}' Z_{jh}' \text{tr}(D_{jh} D_{jh}' \Sigma D_{jh}) + Z_{jh} D_{jh}' Z_{jh}' \sigma_h D_{jh} D_{jh}' \sigma_k' + \text{tr}(Z_{jh} D_{jh}' Z_{jh}') \sigma_k' D_{jh} D_{jh}' \sigma_h I$$

$$(B12) \quad E\{u_h u_k' Z_{ij} Z_{ij}' D_{mn} Z_n' D_{mn}' \Sigma\} = \sigma_{hkc} Z_{ij} Z_{ij}' Z_n' Z_n D_{mn} D_{mn}' \Sigma + \\ + Z_{ij} Z_{ij}' Z_n' Z_n D_{mn} D_{mn}' \sigma_h \sigma_k' + \text{tr}(Z_{ij} Z_{ij}' Z_n' Z_n D_{mn} D_{mn}' \sigma_k \sigma_h')$$

$$(B13) \quad E\{u_h' Z_{ij} Z_{ij}' u_h u_k' Z_m D_{mn} D_{mn}' \Sigma\} = \sigma_h \sigma_k' D_{mn} Z_m' Z_m \text{tr}(Z_{ij} Z_{ij}' Z_n' Z_n) + \\ + \sigma_k \sigma_h' D_{mn} Z_m' Z_m Z_{ij} Z_{ij}' + \sigma_{hkc} \Sigma D_{mn} Z_m' Z_m Z_{ij} Z_{ij}' Z_n'$$

$$(B14) \quad E\{u_h' Z_{ij} Z_{ij}' u_h u_k' Z_m D_{mn} Z_n' D_{mn}' \Sigma\} = \sigma_h \sigma_k' \text{tr}(Z_{ij} Z_{ij}' Z_n' Z_n) \text{tr}(Z_m Z_m' Z_n' Z_n) + \\ + \sigma_k \sigma_h' \text{tr}(Z_{ij} Z_{ij}' Z_n' Z_n Z_m Z_m') + \sigma_{hkc} \Sigma \text{tr}(Z_{ij} Z_{ij}' Z_n' Z_m Z_m')$$

$$(B15) \quad E\{D_{mn} Z_{ij} Z_{ij}' u_h u_k' Z_m D_{mn} D_{mn}' \Sigma\} = Z_{ij} Z_{ij}' D_{mn}' \sigma_h \sigma_k' D_{mn} Z_m' Z_m + \\ + Z_m Z_m' D_{mn}' \sigma_h \sigma_k' D_{mn} Z_{ij} Z_{ij}' + \sigma_{hkc} \text{tr}(Z_m Z_m' D_{mn}' \Sigma D_{mn} Z_{ij} Z_{ij}') I$$

$$(B16) \quad E\{D_{mn} Z_{ij} Z_{ij}' u_h u_k' Z_m D_{mn} Z_n' D_{mn}' \Sigma\} = \text{tr}(Z_m Z_m' Z_n' Z_n) Z_{ij} Z_{ij}' D_{mn}' \sigma_h \sigma_k' + \\ + Z_m Z_m' Z_n' Z_n Z_{ij} Z_{ij}' D_{mn}' \sigma_k \sigma_h' + \sigma_{hkc} Z_n Z_m' Z_m Z_{ij} Z_{ij}' D_{mn}' \Sigma$$

Die Berechnung dieser 16 Erwartungswerte erfolgt nach den Prinzipien, wie sie in Nagar(1959¹) und in Nagar/Gupta(1970)²⁾ angegeben sind.

Beweis von A:

$$(A1) \quad E\{u_h D_{hj} Z_{ij}' u_k\} = E\{D_{hj} u_h Z_{ij}' u_k\} = \\ = (E\{u_t' D_{hj} Z_{ij}' u_k\})_t, \text{ wo } u_t \text{ der } t\text{-te Zeilenvektor aus } U \text{ ist.} \\ = (E\{\text{tr}(u_t' D_{hj} Z_{ij}' u_k)\})_t = (\text{tr}(D_{hj} Z_{ij}' E\{u_k u_t'\}))_t, \text{ wobei}$$

$$E\{u_k u_t'\} \text{ eine } T \times T \text{ Matrix der Form } \Sigma_{tk} = \begin{pmatrix} 0 & \dots & 0 \\ \sigma_{k1} & \dots & \sigma_{kT} \\ 0 & \dots & 0 \end{pmatrix} \leftarrow t\text{-te Zeile}$$

ist, wegen der Voraussetzungen über die u 's.

$$= (\text{tr}(\Sigma_{tk} D_{hj} Z_{ij}'))_t = (\sigma_k' D_{hj} Z_{ij}')_t = (\sigma_k' D_{hj} Z_{ij}')' = \\ = Z_{ij} Z_{ij}' D_{hj}' \sigma_k \quad \text{q.e.d.}$$

1) Nagar, A.L.: The Bias and Moment Matrix..., a.a.O. Appendix.

2) Nagar, A.L. und Y.P. Gupta: a.a.O. Appendix.

$$\begin{aligned}
 (A2) \quad & E\{V_i^* Z_h Q_{h,j} Z_j^* u_k\} = D_i^* E\{U^* Z_h Q_{h,j} Z_j^* u_k\} = \\
 & = D_i^* \left(E\{u_n^* Z_h Q_{h,j} Z_j^* u_k\} \right)_n = D_i^* \left(E\{tr(Z_h Q_{h,j} Z_j^* u_n u_n^*)\} \right)_n = \\
 & = D_i^* \left(tr(Z_h Q_{h,j} Z_j^* E\{u_n u_n^*\}) \right)_n = D_i^* \left(tr(Z_h Q_{h,j} Z_j^* \Sigma_{nn} I) \right)_n = \\
 & = D_i^* \sigma_k tr(Z_h Q_{h,j} Z_j^*) \quad \text{q.e.d.}
 \end{aligned}$$

Beweis von B:

$$(B3) \quad E\{U^* N_{ih} u_h u_k^* N_{jk} U\} = \left(E\{u_l^* N_{ih} u_h u_k^* N_{jk} u_m\} \right)_{l,m} \quad \text{und}$$

$$\begin{aligned}
 & E\{u_l^* N_{ih} u_h u_k^* N_{jk} u_m\} = \\
 & = \sum_{t, t', t'', t'''} E\{u_l(t) n_{ih}(t, t') u_h(t') u_k(t'') n_{jk}(t'', t''') u_m(t''')\},
 \end{aligned}$$

wo $u_l(t)$ das (t, l) -te Element von U und $n_{ih}(t, t')$ das (t, t') -te Element der TxF Matrix N_{ih} ist.

$$= \sum_{t, t', t'', t'''} n_{ih}(t, t') n_{jk}(t'', t''') E\{u_l(t) u_h(t') u_k(t'') u_m(t''')\}$$

Obiger Erwartungswert ist nun genau dann $\neq 0$, wenn die t 's paarweise gleich sind oder wenn alle t 's gleich sind. Das folgt aus den Voraussetzungen, die über die u 's getroffen wurden (vgl. Nagar/Gupta(1970), Appendix). Folgende Fälle sind also zu unterscheiden:

Fall	Erwartungswert
$t=t', t''=t''', t \neq t''$	$\sigma_{lh} \sigma_{km} \sum_{\substack{t, t'' \\ t \neq t''}} n_{ih}(t, t) n_{jk}(t'', t'')$
$t=t'', t'=t''', t \neq t'$	$\sigma_{lk} \sigma_{hm} \sum_{\substack{t, t' \\ t \neq t'}} n_{ih}(t, t') n_{jk}(t, t')$
$t=t''', t'=t'', t \neq t'$	$\sigma_{lm} \sigma_{hk} \sum_{\substack{t, t' \\ t \neq t'}} n_{ih}(t, t') n_{jk}(t', t)$
$t=t'=t''=t'''$	$(\sigma_{lh} \sigma_{km} + \sigma_{lk} \sigma_{hm} + \sigma_{lm} \sigma_{hk}) \sum_t n_{ih}(t, t) n_{jk}(t, t)$

Zusammenfassung der vier Fälle ergibt:

$$\begin{aligned}
 & (\sigma_{1k}\sigma_{hm} + \sigma_{1m}\sigma_{hk}) \sum_{t, t'} n_{ih}(t, t') n_{jlk}(t', t) + \\
 & + \sigma_{1h}\sigma_{km} \sum_{t, t'} n_{ih}(t, t) n_{jlk}(t', t') - \\
 & - (\sigma_{1k}\sigma_{hm} + \sigma_{1m}\sigma_{hk}) \text{tr}(N_{ih} N_{jlk}) + \sigma_{1h}\sigma_{km} \text{tr}(N_{ih}) \text{tr}(N_{jlk})
 \end{aligned}$$

$$\begin{aligned}
 \text{Es ist aber } \text{tr}(N_{ih} N_{jlk}) &= \text{tr}((N^* - (\mathfrak{x}_{ih}/T)I)(N^* - (\mathfrak{x}_{jlk}/T)I)) = \\
 &= \text{tr}(N^* - (\mathfrak{x}_{ih}/T)N^* - (\mathfrak{x}_{jlk}/T)N^* + (\mathfrak{x}_{ih}\mathfrak{x}_{jlk}/T^2)I) = \\
 &= R + \text{Terme von kleinerer Ordnung als } O(1), \text{ die vernachlässigt werden können.}
 \end{aligned}$$

$$\begin{aligned}
 \text{Weiter ist } \text{tr}(N_{ih}) \text{tr}(N_{jlk}) &= \text{tr}(N^* - (\mathfrak{x}_{ih}/T)I) \text{tr}(N^* - (\mathfrak{x}_{jlk}/T)I) = \\
 &= (R - \mathfrak{x}_{ih})(R - \mathfrak{x}_{jlk}).
 \end{aligned}$$

Insgesamt ist also der Erwartungswert des typischen Elements der Matrix in (B3) gleich:

$$(\sigma_{1k}\sigma_{hm} + \sigma_{1m}\sigma_{hk})R + \sigma_{1h}\sigma_{km}(R - \mathfrak{x}_{jh})(R - \mathfrak{x}_{jlk});$$

so daß die gesamte Matrix gleich ist:

$$R(\sigma_{1k}\sigma_h^1 + \sigma_{hk}\sigma_1^1) + (R - \mathfrak{x}_{jh})(R - \mathfrak{x}_{jlk})\sigma_h^1\sigma_{1k}^1 \quad \text{q.e.d.}$$

Die Berechnung der übrigen Erwartungswerte erfolgt nun nach genau dem gleichen Schema wie bei (B3). Der Beweis von (B4) bis (B16) wird demgemäß im folgenden nur kurz dargestellt:

$$(B4) \quad E\{\eta_h u_k' Z_{j1} D_1^{1111} \eta_m \eta_q\} = (E\{\eta_h(t) u_k'(t) \eta_m(t) u_q'(t)\})_{t,q}$$

$$\text{mit } A := Z_{j1} D_1^{1111}$$

Das typische Element ist gleich

$$\sum_{n, t', t'', t'''} a(t', n) \eta_m(t'', t''') E\{\eta_h(t) u_k'(t') \eta_n(t'') u_q'(t''')\}$$

Fallunterscheidung:

$$\begin{aligned}
 t=t', t'=t'', t \neq t'' & \sum_{\substack{n, t' \\ t \neq t''}} a(t, n) n_{\gamma_m}(t'', t'') \sigma_{hk} \sigma_{nq} \\
 t=t'', t'=t'', t \neq t' & \sum_{\substack{n, t' \\ t \neq t'}} a(t', n) n_{\gamma_m}(t, t') \sigma_{hn} \sigma_{kq} \\
 t=t'', t'=t'', t \neq t' & \sum_{\substack{n, t' \\ t \neq t'}} a(t', n) n_{\gamma_m}(t', t) \sigma_{hq} \sigma_{kn} \\
 t=t' \neq t''=t'' & \sum_n a(t, n) n_{\gamma_m}(t, t) (\sigma_{hk} \sigma_{nq} + \sigma_{hn} \sigma_{kq} + \sigma_{hq} \sigma_{kn})
 \end{aligned}$$

Insgesamt:

$$\begin{aligned}
 & \left(\sum_{n, t'} n_{\gamma_m}(t, t') a(t', n) (\sigma_{hn} \sigma_{kq} + \sigma_{hq} \sigma_{kn}) + \sum_{n, t'} n_{\gamma_m}(t'', t'') a(t, n) \sigma_{nq} \right) \\
 & = N_{\gamma_m} \Lambda (\sigma_h \sigma_k' + \sigma_k \sigma_h') + \text{tr}(N_{\gamma_m}) \Lambda \sum \sigma_{hk} \\
 & = Z_j Q_{jl} D_l^i (\sigma_h \sigma_k' + \sigma_k \sigma_h' + (R - \alpha_{\gamma_m}) \sigma_{hk} \Sigma) \text{ bis zur } O(1) \quad \text{a. o. d.}
 \end{aligned}$$

$$(B5) \quad E \{ u_h u_k' u_{jl}'' D_{jl}'' D_l'' \} = \left(E \{ u_h(t) u_k'(t) u_n(t'') u_n(t'') \} \right)_{t, t''}$$

mit $\Lambda = D_j Q_{jl} D_l^i$

Das typische Element ist gleich:

$$\sum_{t', t'', n, n'} n_{jl} (t', t'') a(n, n') E \{ u_h(t) u_k'(t) u_n(t'') u_n(t'') \}$$

Fallunterscheidung:

$$\begin{aligned}
 t=t', t'=t'', t \neq t'' & \sum_{n, n'} n_{jl} (t, t'') a(n, n') \sigma_{hk} \sigma_{nn'} \quad t \neq t'' \\
 t=t'', t'=t'', t \neq t' & \sum_{n, n'} n_{jl} (t'', t) a(n, n') \sigma_{hn} \sigma_{kn'} \quad t \neq t' \\
 t=t'', t'=t'', t \neq t' & \sum_{t', n, n'} n_{jl} (t', t') a(n, n') \sigma_{hn} \sigma_{kn'} \quad t=t' \\
 t=t' \neq t''=t'' & \sum_{n, n'} n_{jl} (t, t) a(n, n') (\sigma_{hk} \sigma_{nn'} + \sigma_{hn} \sigma_{kn'} + \sigma_{hn} \sigma_{kn'})
 \end{aligned}$$

Die TTT Matrix in (B5) hat also die Elemente:

$$\sum_{n, n'} n_{jkr}(t, t''') a(n, n') (\sigma_{hr} \sigma_{nn'} + \sigma_{ln} \sigma_{kn'}) \quad \text{für } t \neq t'''$$

$$\sum_{n, n'} a(n, n') n_{jkr}(t, t) (\sigma_{hr} \sigma_{nn'} + \sigma_{hn} \sigma_{krn'}) +$$

$$+ \sum_{n, n', t'} n_{jkr}(t', t') a(n, n') \sigma_{hn} \sigma_{krn} \quad \text{für } t = t''',$$

d.h. für die Diagonalelemente

$$\text{oder } n_{jkr}(t, t''') (\sigma_{hr} \text{tr}(\Lambda \Sigma) + \sigma_h' \Lambda \sigma_k) \quad \text{für } t \neq t'''$$

$$n_{jkr}(t, t) (\sigma_{hr} \text{tr}(\Lambda \Sigma) + \sigma_h' \Lambda \sigma_k) + \sigma_k' \Lambda \sigma_h \text{tr}(N_{jkr}) \quad \text{für } t = t'''$$

Insgesamt ist (B5) also

$$(R - \infty_{jkr}) \sigma_k' \Lambda \sigma_h I + N_{jkr} (\sigma_{hr} \text{tr}(\Lambda \Sigma) + \sigma_h' \Lambda \sigma_k) \quad \text{oder}$$

$$(R - \infty_{jkr}) \sigma_k' D_{j, j1} D_{j1}' \sigma_h I + \sigma_{hr} \text{tr}(D_{j, j1} D_{j1}' \Sigma) \sigma_k' + \sigma_h' D_{j, j1} D_{j1}' \sigma_k \Sigma \quad \text{zur } 0(1)$$

a.e.d.

$$(B5) \quad E\{u_h u_k' N_{jkr} D_{j, j1} D_{j1}' u\} = (E\{u_h(t) u_k' N_{jkr} u\})_{t, q}$$

$$\text{mit } \Lambda := D_{j, j1} D_{j1}'$$

Das typische Element ist gleich:

$$\sum_{n, t', t'', t'''} n_{jkr}(t', t''') a(n, t''') E\{u_h(t) u_k(t') u_n(t'') u_q(t''')\}$$

Fallunterscheidung:

$$t = t', t'' = t''', t \neq t'' \quad \sum_{\substack{n, t', t'' \\ t \neq t''}} n_{jkr}(t, t''') a(n, t''') \sigma_{hr} \sigma_{nq}$$

$$t = t'', t' = t''', t \neq t' \quad \sum_{\substack{n, t' \\ t \neq t'}} n_{jkr}(t', t) a(n, t') \sigma_{hn} \sigma_{kq}$$

$$t = t''', t' = t'', t \neq t' \quad \sum_{\substack{n, t' \\ t \neq t'}} n_{jkr}(t', t') a(n, t) \sigma_{hq} \sigma_{kn}$$

$$t = t' = t'' = t''' \quad \sum_n n_{jkr}(t, t) a(n, t) (\sigma_{hr} \sigma_{nq} + \sigma_{hn} \sigma_{kq} + \sigma_{hq} \sigma_{kn})$$

Insgesamt:

$$\left(\sum_{n, t'} n_{jlr}(t, t') a(n, t') \sigma_{hlr} \sigma_{nq} + \sum_{n, t'} n_{jlr}(t', t) a(n, t') \sigma_{hn} \sigma_{vq} + \right. \\ \left. + \sum_{n, t'} n_{jlk}(t', t') a(n, t) \sigma_{hq} \sigma_{kn} \right)_{t, q}$$

$$- \sigma_{hlr} N_{jlr} \Lambda' \Sigma + N_{jlk} \Lambda' \sigma_h \sigma_k' + (R - \alpha_{jlr}) \Lambda' \sigma_k \sigma_h'$$

$$+ (R - \alpha_{jlr}) Z_l \gamma_{lj} D_j' \sigma_k \sigma_h' + \sigma_{hlk} Z_l \gamma_{lj} D_j' \Sigma + Z_l \gamma_{lj} D_j' \sigma_h \sigma_k' \quad \text{zur } O(1) \\ \text{a.e.d.}$$

$$(P7) \quad \mathbb{E} \{ u^i N_{ih} u_h' Z_j \gamma_{jl} D_j' u^j \} = \left(\mathbb{E} \{ u^i N_{ih} u_h u_h' A u(t) \} \right)_{q, t}$$

$$\text{mit } A := Z_j \gamma_{jl} D_j'$$

und dem typischen Element

$$\sum_{n, t', t'', t'''} n_{ih}(t', t'') a(t'', n) \mathbb{E} \{ u_q(t') u_h(t'') u_k(t''') u_n(t) \}$$

Fallunterscheidung:

$$t=t', t''=t''', t \neq t' \quad \sum_{\substack{n, t'' \\ t'' \neq t}} n_{ih}(t, t'') a(t'', n) \sigma_{qn} \sigma_{hk}$$

$$t=t', t'=t''', t \neq t' \quad \sum_{\substack{n, t' \\ t' \neq t}} n_{th}(t', t) a(t', n) \sigma_{qk} \sigma_{hn}$$

$$t=t''', t'=t'', t \neq t' \quad \sum_{\substack{n, t' \\ t' \neq t}} n_{ih}(t', t') a(t', n) \sigma_{qh} \sigma_{kn}$$

$$t=t'=t''=t''' \quad \sum_n n_{ih}(t, t) a(t, n) (\sigma_{qn} \sigma_{hk} + \sigma_{qk} \sigma_{hn} + \sigma_{qh} \sigma_{kn})$$

Insgesamt:

$$\left(\sum_{n, t'} n_{ih}(t, t') a(t', n) (\sigma_{qn} \sigma_{hlr} + \sigma_{qk} \sigma_{hn}) + \right.$$

$$\left. + \sum_{n, t'} n_{ih}(t', t') a(t', n) \sigma_{qh} \sigma_{kn} \right)_{q, t}$$

$$- \sigma_{hlr} \sum \Lambda' N_{ih} \sigma_h \sigma_k' \Lambda' N_{ih} + (R - \alpha_{ih}) \sigma_i \sigma_k' \Lambda'$$

$$+ (R - \alpha_{ih}) \sigma_h \sigma_k' D_l \gamma_{lj} Z_j' + \sigma_{hl} \sum D_l \gamma_{lj} Z_j' + \sigma_k \sigma_h' D_l \gamma_{lj} Z_j' \quad \text{zur } O(1) \\ \text{a.e.d.}$$

$$(B2) \quad E\{u_h^{(1)} u_k^{(1)} Z_j Z_{j1} Z_l^{(1)}\}$$

Der Beweis von (B3) verläuft offensichtlich völlig analog zu (B2), wobei man nur N_{jk} durch $Z_j Z_{j1} Z_l^{(1)}$ zu ersetzen hat.

$$(B4) \quad E\{u_h^{(1)} u_k^{(1)} Z_j Z_{j1} D_l^{(1)} Z_m Z_{m1} D_n^{(1)}\} = (E\{u_h(t) u_k^{(1)} B_n(t''')\})_{t, t'''}$$

$$\text{mit } A := Z_j Z_{j1} D_l^{(1)} \text{ und } B := Z_m Z_{m1} D_n^{(1)}$$

und dem typischen Element

$$\sum_{n, n', t', t''} a(t', n) b(t'', n') E\{u_h(t) u_k(t') u_n(t'') u_{n'}(t''')\}$$

Fallunterscheidung:

$$t=t', t''=t''', t \neq t'' \quad \sum_{n, n'} a(t, n) b(t'', n') \sigma_{hk} \sigma_{nn}, \text{ für } t \neq t''$$

$$t=t'', t'=t''', t \neq t' \quad \sum_{n, n'} a(t''', n) b(t, n') \sigma_{hn} \sigma_{kn}, \text{ für } t \neq t''$$

$$t=t''', t'=t'', t \neq t' \quad \sum_{\substack{n, n', t' \\ t' \neq t}} a(t', n) b(t', n') \sigma_{hn} \sigma_{kn} \text{ f. } t=t''$$

$$t=t'=t''=t''' \quad \sum_{n, n'} a(t, n) b(t, n') (\sigma_{hk} \sigma_{nn} + \sigma_{hn} \sigma_{kn} + \sigma_{hn} \sigma_{kn})$$

Zusammengefaßt:

$$\sigma_{hk} A \Sigma B' + B \sigma_k \sigma_h A' + \sigma_h' B' A \sigma_k \quad \text{q.e.d.}$$

(B10) - (B14): Man sieht leicht, daß die Herleitung dieser Erwartungswerte in völliger Analogie zu (B4) - (B8) erfolgt, wobei jeweils nur eine Matrix ausgetauscht werden muß. Auf eine nochmalige Herleitung kann daher hier verzichtet werden.

$$(B15) \quad E\{u_h D_l Z_j Z_{j1} u_k^{(1)} Z_m Z_{m1} D_n^{(1)}\} = (E\{u_h(t) u_k^{(1)} B_n(t''')\})_{t, t'''}$$

$$\text{mit } A := D_l Z_j Z_{j1} Z_l^{(1)}, \quad B := Z_m Z_{m1} D_n^{(1)}$$

und dem typischen Element

$$\sum_{n, n', t', t''} a(n, t') b(t'', n') E \xi_{n'}(t) u_n(t') u_{n'}(t'') u_{n''}(t''') \}$$

Fallunterscheidung:

$$t=t', t''=t''', t \neq t'' \quad \sum_{n, n'} a(n, t) b(t'', n') \sigma_{nh} \sigma_{kn} \quad \text{für } t \neq t''$$

$$t=t'', t'=t''', t \neq t' \quad \sum_{n, n'} a(n, t'') b(t, n') \sigma_{nh} \sigma_{hn} \quad \text{für } t \neq t''$$

$$t=t''', t'=t'', t \neq t' \quad \sum_{\substack{n, n', t' \\ t' \neq t}} a(n, t') b(t', n') \sigma_{nh} \sigma_{h'n'} \quad \text{für } t=t''$$

$$t=t'=t''=t''' \quad \sum_{n, n'} a(n, t) b(t, n') (\sigma_{nh} \sigma_{kn} + \sigma_{nk} \sigma_{hn} + \sigma_{nn} \sigma_{h'n'})$$

Insgesamt:

$$A' \sigma_h \sigma_k' B' + B \sigma_h \sigma_k' A + \sigma_{h'n'} \text{tr}(B \Sigma A) I \quad \text{q.e.d.}$$

$$(16) \quad E \xi_{n'}(t) u_n(t') u_{n'}(t'') u_{n''}(t''') = \langle E \xi_{n'}(t) u_n(t') u_{n'}(t'') u_{n''}(t''') \rangle_{t, r}$$

$$\text{mit } A := u_1 u_1' u_1', \quad B := u_m u_m u_m'$$

und dem typischen Element

$$\sum_{n, t', t'', t'''} a(n, t') b(t'', t''') E \xi_{n'}(t) u_n(t') u_{n'}(t'') u_{n''}(t''') \}$$

Fallunterscheidung:

$$t=t', t''=t''', t \neq t' \quad \sum_{\substack{n, t' \\ t' \neq t}} a(n, t) b(t'', t'') \sigma_{nh} \sigma_{kr}$$

$$t=t'', t'=t''', t \neq t' \quad \sum_{\substack{n, t' \\ t' \neq t}} a(n, t') b(t, t') \sigma_{nh} \sigma_{hr}$$

$$t=t''', t'=t'', t \neq t' \quad \sum_{\substack{t', n \\ t' \neq t}} a(n, t') b(t', t) \sigma_{nh} \sigma_{h'n'}$$

$$t=t'=t''=t''' \quad \sum_n a(n, t) b(t, t) (\sigma_{nh} \sigma_{kr} + \sigma_{nk} \sigma_{hr} + \sigma_{nn} \sigma_{h'n'})$$

$$\text{Zusammengefaßt: } \text{tr}(B) A' \sigma_h \sigma_k' + B A' \sigma_k \sigma_h' + \sigma_{h'n'} B' A' \Sigma \quad \text{q.e.d.}$$

Damit sind die Erwartungswerte (A1) - (B16) vollständig hergeleitet.

LITERATURVERZEICHNIS

Abkürzungen der Zeitschriften:

E	Econometrica
IER	International Economic Review
JE	Journal of Econometrics
JASA	Journal of the American Statistical Association

AITKEN, A.C.: On Least-Squares and Linear Combination of Observation, in: Proceedings of the Royal Society of Edinburgh, vol.55, 1935, 42-48.

ANDERSON, T.W. und H. RUBIN: Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations, Annals of Mathematical Statistics 20, 1949, 46-63.

ANDERSON, T.W.: Estimation of the Parameters of a Single Equation by the Limited-Information Maximum-Likelihood Method, in: Cowles Commission Monograph 10, 1950, 311-322.

ANDERSON, T.W. und T. SAWA: Distribution of Estimates of Coefficients of a Single Equation in a Simultaneous System and their Asymptotic Expansions, E 41, 1973, 685-716.

BASMANN, R.L.: A Generalized Classical Method of Linear Estimation of Coefficients in a Structural Equation, E 25, 1957, 77-83.

BASMANN, R.L.: A Note on the Exact Finite Samples Frequency Functions of Generalized Classical Linear Estimators in Two Leading Over-Identified Cases, JASA 56, 1961, 619-636.

BASMANN, R.L.: A Note on the Exact Finite Sample Frequency Functions of Generalized Classical Linear Estimators in a Leading Three Equation Case, JASA 58, 1963, 161-171.

BASMANN, R.L.: Remarks Concerning the Application of Exact Finite Sample Distribution Functions of GCL Estimators in Econometric Statistical Influence, JASA 58, 1963, 943-976.

- BERGSTROM, A.R.: The Exact Sampling Distribution of Least Squares and Maximum Likelihood Estimators of the Marginal Propensity to Consume, E 30, 1962, 480-490.
- BROWN, T.M.: Simultaneous Least Squares: A Distribution Free Method of Equation System Structure Estimation, IER 1, 1960, 173-191.
- BROWN, G.F., KADANE, J.B. und J.G. RAMAGE: The Asymptotic Bias and Mean-Squared Error of Double k-Class Estimators When the Disturbances are Small, IER 15, 1974, 667-679.
- CHOW, G.C.: A Comparison of Alternative Estimators for Simultaneous Equations, E 32, 1964, 532-553.
- CHOW, G.C.: Two Methods of Computing Full Information Maximum Likelihood Estimates in Simultaneous Stochastic Equations, IER 9, 1968, 100-112.
- CHRIST, C.F.: Econometric Models and Methods, New York 1966.
- DHRYMES, P.J.: An Identity Between Double k-Class and Two Stage Least Squares, IER 10, 1969, 114-116.
- DHRYMES, P.J.: Econometrics. Statistical Foundations and Applications, New York/Evanston/London 1970.
- DHRYMES, P.J.: Distributed Lags, San Francisco 1971.
- DHRYMES, P.J.: Small Sample and Asymptotic Relations Between Maximum Likelihood and Three Stage Least Squares Estimators, E 41, 1973, 357-364.
- FISK, P.R.: Stochastically Dependent Equations, No.21 of Griffin's Statistical Monographs and Courses, London 1967.
- FRIEDRICH, D.: Schätzprobleme sowie Verfahren und Möglichkeiten der Analyse ökonometrischer Modelle, Tübingen 1974.
- GOLDBERGER, A.S.: Econometric Theory, New York 1964.
- GRÜBNER, W.: Matrizenrechnung, Mannheim 1966.
- HAAVELMO, T.: The Statistical Implications of a System of Simultaneous Equations, E 11, 1943, 1-12.

- HAUSMANN, J.A.: An Instrumental Variable Approach to Full Information Estimators for Linear and Certain Nonlinear Econometric Models, E43, 1975, 727-738.
- HENDRY, D.F.: The Structure of Simultaneous Equations Estimators, JE 4, 1976, 51-88.
- JOHNSTON, J.: Econometric Methods, New York 1963.
- KABE, D.G.: A Note on the Exact Distributions of the GCL Estimators in Two Leading Overidentified Cases, JASA 58, 1963, 535-537.
- KABE, D.G.: A Note on the Exact Distributions of the GCL Estimators in a Leading Three-Equation Case, JASA 59, 1964, 881-894.
- KADANE, J.B.: Comparison of k-Class Estimators When the Disturbances are Small, E 39, 1971, 723-737.
- KELLER, W.J.: A New Class of Limited-Information Estimators for Simultaneous Equation Systems, JE 3, 1975, 71-92.
- KLOEK, T. und L.B.M. MENNES: Simultaneous Equations Estimation Based on Principal Components of Predetermined Variables, E 28, 1960, 45-61.
- KOOPMANS, T.C., H. RUBIN und R.B. LEIPNIK: Measuring the Equation Systems of Dynamic Economics, in Cowles Commission Monograph 10, 1950, 53-237.
- KOOPMANS, T.C. und W.C. HOOD: The Estimation of Simultaneous Linear Economic Relationships, in: Cowles Commission Monograph 14, 1953, 112-199.
- LÜDEKE, D.: Schätzprobleme in der Ökonometrie. Identifikation und punktuelle Parameter- und Prognoseschätzung bei interdependenten Systemen. Würzburg/Wien 1964.
- LÜDEKE, D.: Ein ökonometrisches Vierteljahresmodell für die Bundesrepublik Deutschland, Tübingen 1969.
- LÜDEKE, D.: Strukturveränderungen des Investitionsverhaltens im Konjunkturverlauf, Jahrbücher für Nationalökonomie und Statistik, Bd. 188, 1973, 289-319.
- MADANSKY, A.: On the Efficiency of Three-Stage Least-Squares Estimation, E 32, 1964, 51-56.

- MALINVAUD, E.: Statistical Methods of Econometrics, Amsterdam 1966.
- MARIANO, R.S. und T. SAWA: The Exact Finite-Sample Distribution of the Limited-Information Maximum Likelihood Estimator in the Case of Two Included Endogeneous Variables, JASA 67, 1972, 159-163.
- MARIANO, R.S.: The Existence of Moments of the Ordinary Least Squares and Two-Stage Least Squares Estimators, E 40, 1972, 643-652.
- MARIANO, R.S.: Approximations to the Distribution Functions of the Ordinary Least Squares and Two-Stage Least Squares Estimators in the Case of Two Included Endogeneous Variables, E41, 1973, 67-77.
- MARIANO, R.S.: Approximations to the Distribution Functions of Theil's k-Class Estimators, E 41, 1973, 715-721.
- MARIANO, R.S.: Some Large-Concentration-Parameter Asymptotics for the k-Class Estimators, JE 3, 1975, 171-177.
- NAGAR, A.L.: The Bias and Moment Matrix of the General k-Class Estimators of the Parameters in Simultaneous Equations, E 27, 1959, 575-595.
- NAGAR, A.L.: Double k-Class Estimators of Parameters in Simultaneous Equations and Their Small Sample Properties, IER 3, 1962, 168-188.
- NAGAR, A.L. und Y.P. GUPTA: The Moment Matrix of the Two-Stage Least-Squares Estimators of Coefficients in Different Equations of a Complete System of Simultaneous Equations, E 38, 1970, 39-49.
- QUANDT, R.: On Certain Small Sample Properties of k-Class Estimators, IER 6, 1965, 92-104.
- RAO, C.R. und S.K. MITRA: Generalized Inverse of Matrices and Its Applications, New York 1971.
- RICHARDSON, D.H.: The Exact Distribution of a Structural Coefficient Estimator, JASA 63, 1968, 1214-1226.
- ROY, A.R. und V.K. SRIVASTAVA: Generalized Double k-Class Estimators, Journal of the Indian Statistical Association 4, 1966, 38-46.

- SARGAN, J.D.: The Estimation of Economic Relationships Using Instrumental Variables, E 26, 1958, 393-415.
- SARGAN, J.D.: Three-Stage Least Squares and Full Maximum Likelihood Estimates, E 32, 1964, 77-81.
- SARGAN, J.D.: The Validity of Nagar's Expansion for the Moments of Econometric Estimators, E 42, 1974, 169-176.
- SARGAN, J.D. und W.M. MIKHAIL: A General Approximation to the Distribution of Instrumental Variables Estimates, E 39, 1971, 131-169.
- SAVIN, N.E.: Systems k-Class Estimators, E 41, 1973, 1125-1136.
- SAWA, T.: Finite-Sample Properties of the k-Class Estimators, E 40, 1972, 653-679.
- SAWA, T.: The Exact Sampling Distribution of Ordinary Least Squares and Two Stage Least Squares Estimators, JASA 64, 1969, 923-937.
- SCHARF, W.: K-Matrix-Class Estimators and the Full-Information Maximum Likelihood Estimator as a Special Case, JE 4, 1976, 41-50.
- SCHNEERWEISS, H.: Ökonometrie, Würzburg/Wien 1971.
- SCHÖNFELD, P.: Methoden der Ökonometrie, Bd.I, Berlin/Frankfurt 1969, Bd.II, München 1971.
- SCHÖNFELD, P.: A Usefull Central-Limit Theorem for m-Dependent Variables, Metrika 17, 1971, Heft 1-2, S.116-128.
- SRIVASTAVA, V.K.: Three Stage Least Squares and Generalized Double k-Class Estimators: A Mathematical Relationship, IER 12, 1971, 312-316.
- TAKEMUCHI, K.: Exact Sampling Moments of the Ordinary Least Squares, Instrumental Variable, and Two-Stage Least Squares Estimators, IER 11, 1970, 1-12.
- THEIL, H.: Economic Forecasts and Policy, Amsterdam, 1st ed. 1958, 2nd ed. 1961.
- THEIL, H.: Principles of Econometrics, Amsterdam/New York 1971.

- THEIL, H.: Eine einfache Modifikation der zweistufigen Methode der kleinsten Quadrate für unterdimensionierte Stichproben, in: IFO-Studien, Zeitschrift des IFO-Instituts für Wirtschaftsforschung, Berlin/München 1972, S.219-238.
- ULLAH, A. und A.L. NAGAR: The Exact Mean of the Two-Stage Least Squares Estimator of the Structural Parameters in an Equation Having Three Endogeneous Variables, E 42, 1974, 749-758.
- WILKS, S.S.: Mathematical Statistics, New York 1962.
- ZEILNER, A.: An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias, JASA 57, 1962, 348-368.
- ZEILNER, A.: An Introduction to Bayesian Inference in Econometrics, New York 1971.
- ZEILNER, A. und H. THEIL: Three-Stage Least-Squares: Simultaneous Estimation of Simultaneous Equations, E 30, 1962, 54-78.

LEBENS LAUF

- 12.4.1946 geboren in Mannheim als Sohn des
Oberingenieurs Willi Scharf und
seiner Ehefrau Luzie, geb. Elzer
- 1952 - 1956 Besuch der Volksschule in Mannheim
- 1956 - 1965 Besuch des Tulla-Gymnasiums in Mann-
heim
- 19.2.1965 Abitur
- 1.4. - 20.9.1965 Industriepraktikum
- 1965 - 1968 Studium der Mathematik in Berlin
- 1968 - 1973 Studium der Mathematik und der Wirt-
schaftswissenschaften in Freiburg
- 16.2. - 15.4.1970 Werkstudent am Statistischen Bundes-
amt in Wiesbaden
- 4.7.1973 Diplom in Mathematik
- 1.11.1973 - 31.3.1977 wissenschaftlicher Mitarbeiter
am Institut für Allgemeine Wirtschafts-
forschung, Abt. Statistik und Ökono-
metrie, der Universität Freiburg;
Doktorand bei Professor Dr. D. Lüdeke
- 20.1. - 22.1.1975 vortragender Teilnehmer am Winter-
Symposium der Econometric Society
in Genf

